

New research machinery needed

Too many promising international research collaborations, from AIDS research to the sequencing of the human genome, languish for lack of a workable framework for tangible and short-term research.

THAT research is an international enterprise whose long-term benefits are international is a truism nowhere questioned, even where governments seek immediate national advantage from what they spend on research. What has been overlooked in the past few years is that even the most self-interested governments often recognize that some objectives in research can reasonably be sought only by research that is strictly international. Europe's sensible recognition 30 years ago that collaboration was the best way of keeping an interest in high-energy physics has been vindicated by the success of the high-energy physics laboratory, CERN, at Geneva. Much the same is true of the European programme for thermonuclear fusion research, supported by the European Commission, which is represented by the Joint European Torus, running well in rural England. The European Southern Observatory is another example of success. These collaborations have succeeded because their management has been firmly in the hands of technical people. In these cases, success has also been determined by the need for equipment more expensive than collaborators individually could have afforded. Now, there is emerging a rash of projects that cry out for collaboration because they are beyond the intellectual resources of single countries.

The projects stalled for want of a means of organizing international cooperation include that for generating a nucleotide sequence for the human genome, where several research groups and national organizations are keen to go ahead with an enterprise which, if not a worthwhile end in itself, could be of incalculable value in biomedical research in decades hence. Thus the US Department of Energy and the Japanese Science and Technology Agency have an interest in organizing and supporting the project (see p. 195); each seems sensibly to have decided that two independent projects would be a waste of resources and a source of confusion, but to differ sufficiently in their objectives as to impede agreement between themselves, let alone with others. Much the same is true of the growing recognition in Europe that the governments interested in backing programmes to discover means for preventing the spread of AIDS (acquired immune deficiency syndrome) and for the more effective treatment of the disease should do so in collaboration. An even more challenging proposal (from Japan) is that there should be an international programme of research directed at an understanding of the mechanism of high-temperature superconductivity based on the sensible assertion that the prize of room temperature superconductivity will be attainable only when there is an understanding, now lacking, of why variants on the theme of lanthanum-barium-copper oxide should be superconducting above the temperature of liquid nitrogen.

Dilemma

All of these are projects which, if taken up at all, should be done quickly. Yet, as things are, national grant-making agencies cannot tell how much they should be spending. Shortages of people and ideas are the chief reasons for collaboration. Such projects are also largely centred on programmes of basic research which, nevertheless, are likely to yield substantial benefits to commercial companies. But projects of this kind

present even those governments with an interest in a successful outcome with the dilemma that there is no operational justification for them to ask that the funds they commit will be spent by their own nationals in ways that they determine. If there were to be a European AIDS research programme, or even collaboration between Britain and France as the French have suggested (see p. 191), the best way of working is that decisions on who does what should be made not by the moneybags or their officials but by groups of scientists in the field who are able and willing to make judgements on the quality of different people's work.

Foundation

These features of potential collaborative projects are also some of the reasons why they now hang fire. That is why some thought should be given to their common solution. The obvious mechanism is that of what might be called the limited research foundation. Interested national governments or their grant-making agencies should commit funds in agreed proportions to a body with a constitution like that of a charitable foundation operating internationally. Decisions on how the money should be spent would then be taken not by the donors but by the trustees appointed for the purpose, who would also be responsible for telling when and how commercial companies (which may also be donors) should benefit. There are plenty of illustrations of how arrangements like these work well nationally, in the enlightened management of national laboratories as different as the Max-Planck institutes in West Germany and observatories such as Kitt Peak in the United States.

With the growing need of international collaborations, and their inherent virtues, it is high time that such schemes were given a fair trial. Japan's Ministry of International Trade and Industry (MITI), convinced that its Ministry of Finance will not allow it to support a sufficient programme of basic research in high-temperature superconductivity, could find that a programme mounted internationally would strike an encouraging response elsewhere. Why not give that a try? The proposal for collaboration on AIDS research in Europe is more complicated, if only because of the British government's refusal since last December to agree to the European Commission's research proposals for the next five years. (The budget has now been agreed for the present year at last year's level.) But Britain's objections to the Commission's way of carrying out research could well be resolved if the Commission's programme were broken into smaller pieces, with objectives tangible enough to be used in the evaluation of success and precise enough to be managed by technical people (not officials) deciding who should do what without regard to the ambitions of Europe's member states that each should get more spending out of common projects than it contributes financially. The French and British medical research councils, each already embarking on national programmes of AIDS research, should quickly pool resources. It is unlikely that, with the resources at their disposal (say \$15 million a year), they will be able to cure AIDS, but they might at least provide a much-needed example of how international research should be carried out. □



Banks in Queer Street

Western banks should tell their customers they have lost money on loans to poor countries.

FUNNY things are happening in the West's commercial banks. For the past two months, many of the banks have been saying that their assets are worth less than previously thought. The curious consequence has been that those who invest on the world's stock-markets have judged the same banks to be worth more. This phenomenon deserves more attention than it has been given, because of its bearing on the relationship between the rich and the poor countries of the world.

The story begins in 1973, when petroleum-exporting countries (OPEC) increased the price of crude oil by a factor of three, which later grew to a factor of five, and began accumulating surpluses of more than \$100,000 million a year from the revenues they could not spend. The strange blend of inflation and slump in Western economies that followed is, in retrospect, easily explained. Dearer oil meant higher costs, but the transfer of resources to the oil producers meant that industrial states had less cash to invest in their own development. For a time it seemed as if the industrial countries would be in serious trouble, but it then emerged that the oil producers had no choice but to lend back their surpluses to the commercial banks. "Recycling" became the name of the bankers' game. Meanwhile, the poor countries, less attractive havens for oil producers' cash, were unable to pay for the oil they needed to stay afloat. Industrial governments, conscious of a moral pressure to help the less fortunate, then hit on what seemed a neat solution: persuade the commercial banks to recycle part of the oil producers' surpluses to the poor countries. Thus began the process by which the pauperization of the poor was made permanent. The debts of the developing countries, in round numbers a cool \$500,000 million, are growing as interest they cannot afford to pay is added to the principal sums outstanding.

Some exceptions to this general pattern stand out. Among industrial oil-consumers, West Germany and Japan always manfully paid for their oil at its higher price. Britain, apparently exempted from physical dependence on the oil producers by North Sea oil, was nevertheless compelled to pay the higher prices, without which much of the oil production of the past decade would not have been economic. Among the poor countries, India seems to have emerged most successfully, largely because of a green revolution and an increase in its industrial production. But once-proud poor oil-producers such as Mexico, Nigeria and Venezuela, having borrowed recycled money in the hope of sustaining general development are now saddled with debts they cannot pay.

For five years, most governments have been hoping that this mess would go away. Now the commercial banks have begun to acknowledge that it will not. West German and Swiss banks have been for five years prudently providing for the chance that some of their loans would never be repaid, but the big break came in May, when Citicorp, the largest US bank company, set aside \$3,000 million as an earnest of its diminished expectations. Other US banks followed suit, as have two in Britain. The most startling development so far is the decision of the British Midland Bank to sell off a fifth of its banking business in Britain and to drum up from its existing shareholders an extra £700 million, so as to set aside an extra £915 million in respect of overseas loans unlikely to be required. The bank's shareholders have been told that it will cost them an extra £700 million to own a safe and profitable bank four-fifths the size of the business they believed they owned. That the shares of these banks have almost always increased in value is said to be a measure of the general welcome for their realism.

The consequences of these developments will be serious. Having had their fingers burned, the banks may now be tougher on future applications for loans to poor countries, which could

be especially serious when almost all the governments of Latin America are queuing up in New York to renegotiate their present debts. Can it make sense that these negotiations should be left to the commercial banks, with a vested interest in making what they can of the assets whose value has been written off? It would be better that the banks writing off these assets (with the help of their governments' willingness not to tax the sums set aside to cover bad debts, which would otherwise have been profits) should be compelled to sell them at knockdown prices to, say, the International Monetary Fund, which could then negotiate with the debtor countries an equitable liquidation of the loans. But this is merely a crude solution of the problem. The governments and banks that created the imbalances in the first place should now recognize that the responsibility for finding a lasting solution rests with them. □

How to feed worlds

The United States advocates rationality in agriculture, not before time.

THE Reagan administration in the United States, which has not enjoyed much adulation in recent months, deserves high marks for what it had to say about agriculture at Geneva ten days ago. Briefly, US trade representatives helping to plan the intended renegotiation of the General Agreement on Tariffs and Trade (GATT) said that one of their objectives was the abolition of all agricultural subsidies by all governments within a decade. The farmers in the Middle West baulk at the news, but elsewhere there should be an unstinted welcome for this notion.

Why should agricultural subsidies matter so much that mere talk of their removal is good news? Everybody knows that, at least in industrial countries, governments subsidize agriculture. In the old days, when many countries were forced to be self-sufficient economic units, there were respectable excuses. Given climatic fluctuations, it made sense to sustain production at levels that would be sufficient even in bad years. It is also proper that governments should pay some attention to the welfare of farming communities. The costs of these policies appear in one of two ways. Either people must pay higher taxes so that farmers can be subsidized, which is the pattern in the United States, or food prices are manipulated so that food-consumers pay farmers directly, which is the European practice.

The immediate losers are the consumers, who pay more for food than would otherwise be necessary. In most places, consumers do not complain too loudly; farm subsidies tend to be a small fraction of gross national product and country-folk are likeable people. But consumers do from time to time protest; even well-to-do Japan is occasionally up in arms at the high price of rice, sustained at several times the world market price by a system of guaranteed farmers' prices (just reduced by 6 per cent) and import controls. But the big money is spent in Europe and the United States (roughly \$30,000 million a year cash), some of it during the past year in underbidding for the markets for grain in places such as Egypt. What seems now to have happened is that the United States has realized that further intensified competition along these lines is a recipe for bankruptcy. Europe, wedded to the ideology of its Common Agricultural Policy, cannot see the danger.

The more real present losers in this competition are not the principal actors but those who stand on the sidelines — some major food producers (Canada, Argentina, Australia and New Zealand) and many potential food-producers, several of the developing countries in which food production would increase at least to the point of self-sufficiency if the markets were not rigged against them. More generally, agricultural subsidies are a device for making sure that people everywhere work at tasks that suit them less than well. If the United States can make the lesson stick — and make the US Congress listen — it will have done a great public service. □

Wellcome versus Genentech result

- TPA patent case ends with 90-page judgement
- Wellcome's 'victory' may not be decisive

London

THE confusion and controversy surrounding biotechnology patents seems no closer to resolution, following the British High Court decision declaring invalid the wide-ranging patent held by the US firm Genentech on the clot-dissolving agent, tissue plasminogen activator (TPA). Lawyers are still studying the highly complex 90-page decision which upholds the patent challenge by the UK drugs firm Wellcome, and the two sides disagree on exactly what the judge, Mr Justice Whitford, is saying about what he calls "the potential goldmine" of TPA production.

Wellcome has greeted the decision "with satisfaction", saying it vindicates their arguments that Genentech's patent failed to meet the criteria of novelty and inventiveness. Genentech lawyer Simon Cooke disagrees with that interpretation, and says the decision shows the judge believes the product is in fact novel and inventive. And Genentech's chief executive officer, Robert Swanson, says the company will appeal the decision rather than narrow its patent claims.

The decision includes a discussion of each of the 20 claims covered by Genentech's patent, ranging from processes for making TPA to pharmaceutical compositions containing it. One claim covers "human TPA as produced by recombinant DNA technology", and another covers "all processes for producing TPA by recombinant DNA technology."

The decision accepts that Genentech was the first to discover the full sequence for producing TPA, and as such "there might have been scope for a limited process claim"; but that discovery "does not justify the broad claims made." The claim covering any DNA technology for TPA "is too wide and is bad. There is no basis for it," the decision states.

The judge says, "What Genentech did was achieved by what, in my judgement, was rather more than the exercise of proficiency — it involved laborious and costly effort, and to deny any monopoly protection to those who are prepared to put as much time, skill and money into research as Genentech did is only too likely to discourage workers in this field from making advances which may be of the greatest public benefit." But in what might be viewed as a contradictory statement, he continues, "To grant them a monopoly which would stop others attempting to discover alternative, possibly wholly unknown and possibly better routes to that end, would be to stifle research which, in

the public interest, it ought to be open to other investigators to pursue and over which other investigators, in their turn, if they make a valuable contribution, might be able to secure proper protection."

Genentech's lawyer says the judge has "misunderstood the technology", and has made a decision that is "fundamentally wrong". Mr Cooke adds, "Genentech has given to the public the structure of the gene for TPA, and he's in effect said we should get no protection whatsoever".

But Genentech's patent claim is "too greedy", according to London patent attorney Iain Baillie. He predicts that the

company will find itself "in the same mess" if it has filed a similarly wide-ranging claim for a patent in the United States, and he also forecasts a loss for Genentech in the Court of Appeals. Wellcome says if a similar patent is granted in the United States, Genentech could expect to fight another court battle.

It is likely to be at least a year before the British Court of Appeals can hear what has now become a case of legal technicalities, and another year if the legal battle goes on to the House of Lords. Meanwhile, Wellcome is continuing to produce TPA for use in its clinical trials; under US law, this is not considered a patent infringement. The legal battle has done nothing to stop the race for the lucrative TPA marketplace.

Kathy Johnston

... US patents in the balance

Washington

As the first major biotechnology patent to be disputed in the United Kingdom, the TPA case is expected to serve as an indicator of how readily future broad biotech patents will be granted and upheld. But the British opposition to broad claims may not carry over the Atlantic, because the current US biotech patent climate is relatively liberal.

The bad news falls on Genentech just weeks after the United States Food and Drug Administration (FDA) withheld marketing approval for Activase, its version of TPA (see *Nature* 327, 450; 1987). But, of the 19 or so companies that are working on TPA, Genentech still leads the pack in the race for FDA approval. Stuart Weisbrod, a biotechnology stock analyst for Prudential-Bache, says Genentech has a two-year lead on the competition in the race for FDA approval despite the setback, and expects them to gain marketing approval by the end of the year. Wellcome does not expect to be able to sell its version of TPA until 1989. Genentech now has approval to market Activase in France, the Philippines and New Zealand.

Investors were not greatly dissuaded by the outcome of the case. On the day the news broke, Genentech's stock dipped \$1.375, to close at \$38.625 a share. This is in contrast to the day the FDA denied approval of Genentech's TPA, when its stock plummeted \$11.50 per share.

Most observers feel that Genentech will succeed in obtaining a patent that allows them access to at least a piece of the estimated \$1,000 million pie for worldwide TPA sales. But Wellcome, and collaborations between Smith-Kline and Biogen, and Hoechst and Chiron are reported to

be producing TPA by essentially the same process as Genentech, so patent coverage is vitally important. The company that gets the patent for the process can lock the others out, forcing them to develop another means to produce the drug. Using a new process would force those companies to start the FDA approval process all over again.

Genetics Institute is the furthest along in developing a "second generation" form of TPA that may make the current patent dispute over TPA moot. According to vice-president Robert Kamen, Genetics Institute has deleted the fibronectin and epidermal growth factor domains of the sprawling TPA molecule, and have removed three glycosylation sites. The resulting stream-lined version may have a half-life in the bloodstream of up to one hour, compared to 5 to 10 minutes for the natural molecule. This means it can be administered in one bolus injection to break down a clot, instead of requiring several injections, as does the natural version.

Genetics Institute plans to manufacture its second-generation TPA at WelGen, a joint mammalian cell production facility it is building with Wellcome, but Wellcome does not own any rights to Genetics Institute's TPA. Others working on second-generation TPA include Smith-Kline, Integrated Genetics, Eli Lilly, Searle, and Genentech.

Genetics Institute and Integrated Genetics have applied for patents for their modified forms of TPA. Theirs may eventually prove out to be the strongest patent positions, as it is easier to defend a modified molecule as novel than a slight variation on tried-and-true genetic engineering techniques.

Carol Ezzell

DFG aiming high with its latest astronomy wants list

Bonn

MAJOR improvements in West Germany's astronomy programme were urged by the Deutsche Forschungsgemeinschaft (DFG), West Germany's basic research funding organization, on 9 July. Structural problems in the universities and insufficient funding are preventing West German students from becoming top-class astronomers and astrophysicists, says a DFG report. The report calls for additional funding of DM771.5 million over the next 15 years to be invested in a variety of new projects as well as a restructuring of West German astronomy.

But the DFG is not backing up its recommendations with money. Instead, it is calling upon both the federal government and the *Länder* (states) to increase their support of astronomy. The DFG pays out about DM20 million per year for astronomy projects, a figure which it is not planning to increase by more than a few per cent a year.

Recent budget cuts in *Länder* such as Lower Saxony make it doubtful that all the DFG proposals will be realized. But DFG President Hubert Markl points out that West Germany is "not a poor country" and that temporary problems in some *Länder* should not divert attention from the goal of a better programme.

University research in astronomy in West Germany is decentralized, having developed around a number of small optical telescopes based at various universities. But significant basic research is now done mainly on large optical telescopes, radio telescopes and other massive pieces of equipment.

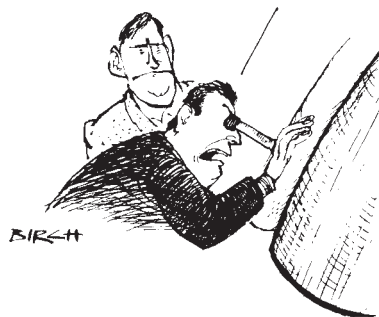
West German universities are poor relations compared to the Max Planck Institutes (MPIs). The institutes involved — the MPIs for Astronomy and Nuclear Physics in Heidelberg, for Radioastronomy in Bonn and for Physics and Astrophysics in Munich — spend significantly less in administrative costs and are at the same time free to concentrate on research. The potential for further decoupling is a "danger", in the view of Gisbert Winnewisser, director of the physics institute at the University of Cologne.

One of the most urgent problems in West German astronomy is the lack of computer hardware and software for image processing. To remedy this shortage, the DFG recommends that a number of VAX computers be installed in universities, all using international compatible software. Such computers have already been installed in single institutes. The report also urges the formation of a "German Astro-Net" for the exchange of data among universities.

The DFG's other recommendations include a more realistic budget for travel for university professors and students as well as more funds for additional focal plane instruments to be designed at individual universities and used on the various large telescopes accessible to West German researchers.

If the funds can be found, West German astronomers and physicists will embark on a variety of new projects in the next few

*I can't see anything for that
dam' butter mountain...*



years. Millimetre and submillimetre very long base line interferometry (VLBI) will be supported through the creation of a European VLBI Computation Centre in the Netherlands capable of synthesizing and evaluating data from up to 16 telescopes. West Germany would provide 25 per cent of the estimated DM30 million

cost. If this project is not funded, the DFG recommends the expansion (at a cost of DM 5 million) of the VLBI Correlator at the MPI for Radio-astronomy in Bonn.

A "gravity-wave telescope", will be built, provided that DM60 million is made available for the project. A laser interferometer is planned with two 3-km-long arms. The United States, the United Kingdom and France are interested in carrying out similar projects.

West German researchers are also interested in joining those investigating extra-solar neutrinos. Heinz Völk of the MPI for Nuclear Physics explained proudly that this was a goal even before the discovery of neutrinos from the supernova in the Large Magellanic Cloud. But Völk would not reveal the type of materials that would be used in the proposed detector, which is to be constructed (if funded) in an abandoned coal mine in West Germany.

Once before, in 1962, the DFG issued a similar report on astronomy in West Germany. Its recommendations were for the most part adopted, with highly favourable results. Despite recent stagnation in support for basic research in universities, Markl thinks the chances are good that much of the DM771.5 million programme will be supported by the Ministry for Research and Technology (BMFT). "The BMFT asked us to hurry up and produce the report so that they could consider our recommendations at their next internal meeting", said Markl. "I think we will be able to realize a lot of what we have asked for."

Steven Dickman

Support for basic research under threat

West Berlin

THE Deutsche Forschungsgemeinschaft (DFG), which funds basic research in West German universities, feels threatened by the budget axe. Its president, Hubert Markl, warned delegates to the organization's annual conference in West Berlin on 7 July that basic research funding may be regarded as an "unnecessary subsidy" by the architects of the forthcoming tax reform.

The DFG budget will rise 3.5 per cent for 1988, a very small increase in the light of the 12 per cent growth in grant money requested by researchers in the past year. By comparison, the Max Planck Society received a 5 per cent increase for its basic research institutes for fiscal year 1988.

Markl pointed out that even the 3.5 per cent increase may be frozen, just as were the increases of the past two years. That would make money inaccessible to the DFG until late in 1988 at best.

Markl called attention to the 17 per cent budget increase proposed for the US National Science Foundation (NSF) for 1988, and the longer-term plans to double

the NSF budget. He also warned of stronger competition from countries such as Italy and Japan over the next few years.

The pressure on the DFG will be especially severe in the next few years because of a 'bulge' in the West German student population. Some of the same young people who are now crowding universities to 150 per cent capacity will soon be applying for research funding. It would be "economic madness", said Markl, to stop new funding for university research now, after DM50 million has been invested in higher education over the past 25 years.

One new programme, initially supporting five people aged under 33 to the tune of DM200,000 a year for 5 years, is designed to encourage outstanding young researchers and prevent them from leaving academic life.

In his address to the Berlin conference, Markl emphasized the need for the West German government to consider its priorities. The sale of subsidized butter to the Soviet Union alone cost almost as much in 1986 as the DFG budget of DM1,029 million, he said.

Steven Dickman

French AIDS research plan launched

Paris

THE appointment of Professor Jean-Paul Levy to coordinate AIDS (acquired immune deficiency syndrome) research augurs well for the success of the French national research programme, announced in May (see *Nature* 328, 3; 1987). Levy (53) is professor of haematology at the University of Paris V, is director of an immunology and oncology laboratory at the Hôpital Cochin in Paris and has worked on cellular immunology and retroviruses since the 1960s.

Part of the strength of Levy's new position is that he will be responsible for coordinating work supported by the two principal research councils, CNRS (Centre National de la Recherche Scientifique) and INSERM (Institut National de la Santé), which at present support nearly all AIDS research (including some at the Institut Pasteur). The two research councils will spend about 60 per cent of the FF100 million (£10 million) budget allocated by the science ministry to the new research programme, with the remainder channelled through the Institut Pasteur, industrial companies such as Transgène and other smaller groups.

Levy, who has been a member of the INSERM directorate since 1980, says that while the link between CNRS and INSERM is an essential feature of the new research programme, it will nevertheless be subordinate to the Programme National de Recherche sur le SIDA, known as PNRS, whose direction will be determined by the government-appointed committee which includes science minister Jacques Valade (chairman), biochemist Pierre Louisot (vice-chairman) and Alain Pompidou from the Ministry of Health.

Levy believes that he and the PNRS committee have a strategic function. Although at first concerned with the selection and coordination of research projects and the allocation of funds, their objective is "to develop AIDS research and, more generally, research in human retroviruses in France to the greatest possible extent". Although 40 INSERM and 25–30 CNRS laboratories are already involved in AIDS research, Levy suspects that this may not be sufficient "if we are to compete internationally", for which reason he believes part of his task will be to help reorientate certain research groups towards "major problem areas". This may entail the creation of new laboratories, as with the intended move, later this year, of Jean-Claude Chermann from the Institut Pasteur to set up a virology laboratory at the Luminy University (near Marseilles).

In contrast with the programme of the British Medical Research Council

(MRC), whose more limited budget has suggested a concentration on drugs and vaccines (see *Nature* 327, 355; 1987), the French programme will be broadly based. It is now proposed that representatives of laboratories working on 12 major research themes should take part in discussions to decide what work deserves support and, ultimately, which laboratories should do the work. Levy and the PNRS committee will then allocate budgets and other resources and encourage links with industry for the *in vitro* testing of antiviral molecules and other screening tasks.

The final distribution of the funds for 1987–88 will not be decided until all the project groups have met, but priority will be given to the search for a vaccine, trials of antiviral and other therapeutic agents and epidemiology. It is also intended to create new virology and screening laboratories throughout France, recruiting the people needed from abroad if necessary.

Research programmes whose organization has already been decided include an attack on the nucleoside inhibition of reverse transcriptase, the inhibition of the TAT protein of the human immunodeficiency virus (HIV), nonsense RNA,



Professor Jean-Paul Levy — "Monsieur AIDS". studies of the viral coating and the transmission of virus from mother to fetus.

According to Louisot, the new AIDS budget, which is an addition to existing research funds, is a "significant" investment by the French government and should benefit "virology and molecular biology as a whole". Saying that "for once we have an initiative that is intelligent, well conceived and constructive", Louisot also says that the budget will have to be maintained for several years if it is to be effective and that, even then, the new budget will be dwarfed by the funds available in the United States. That is why, Louisot says, France will be looking, in August this year, for an enlargement of the research programme on a European scale. He believes that a joint study by INSERM/CNRS and MRC could be a starting-point for such a project.

Peter Coles

Spain's plans for science

Barcelona

THE CSIC (Consejo Superior de Investigaciones Científicas), the Spanish research council, has outlined a programme that, for the first time, defines the activities of its 1,700 researchers and 94 institutes during the next five years. The programme has been presented to the Interministerial Commission of Science and Technology for possible inclusion in the National Plan for Research, now being drafted. The programme (CSP) is a 536-page document containing a description of its purpose and methodology, how research groups should apply for funds and a proposed budget for CSIC. The main part of the document is devoted to the CSIC's scientific objectives which are grouped in 22 subprogrammes.

One of the subprogrammes is called "General Promotion of Knowledge" and will cover basic research not included elsewhere, including relations of Spain with the United States and with the Arab world.

The budget of CSIC will rise steadily during the five-year period to 30,000 million pesetas (£150 million) in 1992, a figure considered conservative in the light of increased expenses in many research fields. It allows for 6.5 per cent annual increase in CSIC's scientific staff to 2,500 researchers in 1992, and around 2,000 fellowships for graduate students and postdoctorals.

P.P.

Indian/Soviet agreement

New Delhi

INDIA and the Soviet Union have become partners in frontier areas of research under a pact on science cooperation signed in Moscow last week by the Indian prime minister, Rajiv Gandhi, and the Soviet leader, Mikhail Gorbachev. The agreement, which covers the period up to the year 2000, was set in train during Gorbachev's visit to New Delhi last November, and a working party outlining the agreement now reached was set up in Moscow in March (see *Nature* 326, 632; 1987). K.S.J.

Bulgaria and the sea

London

BULGARIA, which at the beginning of the year declared its right to a 200-mile offshore "economic zone" along the Black Sea coast, has announced the establishment of a "technological centre for marine resources". Charged with the task of "undertaking national work on studying and utilizing the riches of the sea", the centre is an interdisciplinary association of (so far) 14 governmental, economic and academic bodies. According to the daily *Otechestven Front*, the centre has already started work on six "urgent" projects, including designing a towable submersible with an operational depth of 3,000 m and technology to allow divers to work outside a craft at a depth of 200 m.

V.R.

Seismic monitoring of Soviet nuclear tests to resume

Washington

SEISMIC monitoring of the Soviet nuclear test facility in Kazakhstan by a private US organization will continue for another 14 months following an agreement signed in Moscow last month, but the ground rules are to be altered. Recordings will be permitted to continue during nuclear tests, but the stations will have to move at least 600 miles from the test site, five times further than they are now.

The National Resources Defense Council (NRDC), a private non-profit environmental protection organization, first established the listening posts last summer in conjunction with the Soviet Academy of Sciences (see *Nature* 321, 638; 1986). The posts were intended to show that effective monitoring was no obstacle to a test-ban treaty. But when the stations were first established, the Soviet Union was observing a unilateral moratorium on nuclear testing. When tests resumed early in February, the three stations were required to shut down.

Under the new agreement signed with the Soviet Academy, the stations will be permitted to continue to operate during bomb tests, a move NRDC staff member Jacob Scherr sees as a significant shift in Soviet policy. Part of the reason for the change of heart is the existence of a US-sponsored seismic station at Urumqi in northern China, about 600 miles south of the Kazakh test site. NRDC stations will be located to the north and west of the test site.

The US government has not been comfortable with the NRDC effort. State Department officials say it is not the job of private organizations to conduct US

foreign policy. The United States prefers a monitoring scheme called CORTEX (continuous reflectometry for radius versus time) that requires a bore hole 50 feet from the test chamber. The Soviet Union has shown no interest in this scheme, proposing instead that the countries explode one of their own bombs at the other's test site. International seismic stations could then be calibrated for yield determination starting with a known figure. After an initial flurry of interest, the bomb-swap scheme seems to have died quietly.

The existing monitoring stations will continue to operate for the remainder of this year, but will shut down during bomb tests. The new agreement calls for five stations, the first of which could be established as early as this November. The project's base has been moved from Karkaralinsk, Kazakhstan to Garm, Tadjikistan.

Joseph Palca

Vera Rich adds: According to Soviet sources, the US equipment will be calibrated by means of a non-nuclear (chemical) explosion in the next few weeks. The Soviet side have hailed the extended agreement as a confirmation of Soviet readiness to take part in "diverse verification undertakings", including, possibly, the establishment of an international verification system.

Meanwhile the Soviet oil industry is continuing its series of 20-kilotonne nuclear explosions in the Ural oil fields, events which, the Soviet spokesman say, have nothing in common with bomb tests, except that both kinds of explosions are nuclear and underground. □

Chernobyl defendants in defiant mood

London

THE trial of the six persons accused of responsibility for the Chernobyl disaster opened in Chernobyl town last week. According to indictment, the six, during the course of a number of years, repeatedly violated operating instructions and, in particular, on the night of 25/26 April 1986, whilst conducting experiments on the No. 4 generating set, "crudely violated" operating rules, thereby causing the explosion. The six were indicted under Article 220/2 of the Ukrainian Criminal Code, covering breaches of safety regulations at explosion-prone enterprises.

Although only the first day of the trial was open to the media, it is clear that the accused are prepared to put up a fight. Viktor Bryukhanov, former director of the power station, says he was not on duty at the time of the accident, Mykola Fomin,

former chief engineer, blames the design of the reactor and "technical shortcomings", Anatol' Dyatkov, his former assistant, blames a construction fault, and Boris Roguzhin, head of the shift on duty at the time of the accident, maintains that up to the moment of the accident, he was working according to the rule-book.

Responsibility for the accident is clearly to be confined to the on-site personnel — the others are engineers Aleksandr Kovelenco and Yuri Lanshkin. A subsidiary charge against them is that immediately after the accident they gave distorted information about the radiation situation and did not take the necessary precautionary measures. If this is proved, it will make them formally responsible for the delays in evacuating the affected areas and in notifying neighbouring countries of the danger.

Vera Rich

Locust swarms threatening

London

HUGE swarms of desert locusts in the Red Sea coastal areas of Ethiopia are under attack by a few pesticide-wielding aircraft, in a desperate battle to prevent a major outbreak in northern and western Africa. But there are serious obstacles, including inadequate resources and civil strife.

"We're worried that we will get major swarms over other areas within a few months", says Lukas Brader, director of the Emergency Centre for Locust Operations of the UN Food and Agriculture Organization (FAO) in Rome.

In the past few weeks, a giant cloud of desert locusts covered many square kilometres in Eritrea, including Asmara airport. The locusts, which have found favourable breeding conditions with the return of the rains, have been moving into northern Ethiopia and eastern Sudan in ever-increasing numbers, the worst outbreak in several decades.

A control programme would have faced the best chances of success earlier this year, when the newly-hatched insects could not fly and would have been easier to kill. But the spraying programme in spring "was not as effective as we had hoped", according to the FAO's Jeremy Rossey, who blames the fact that ground teams could not safely work in the strife-torn areas of northern Ethiopia. In addition, locust control organizations have simply not been geared up for a large-scale invasion, after winding down and concentrating on other pests during the prolonged drought when locusts were not a problem. Surveying and forecasting teams were disbanded. "Complacency about locusts really set in", says one specialist.

Now, instead of a coordinated system of control, the situation has become one of "crisis management", some scientists say. But the resources are inadequate to deal with the crisis. Supplies of the pesticide fenitrothion are running low at the Desert Locust Control Organization for East Africa in Addis Ababa, and there are only three aircraft available for spraying work.

The European Economic Community has just donated 50,000 litres of the pesticide, enough for 1,500 km², according to Rossey. Now that the swarms have moved to very rough terrain, however, spraying will be much less effective, or even impossible, as the insects shelter in ravines up to 1,000 metres deep.

The danger is that the locust swarms could move to the farming areas of the African Sahel; already, some crop damage is being reported in northern Ethiopia. The infestation may spread back to the Red Sea coast, or to the Ogaden Desert area.

Kathy Johnston

Cabinet Office bullish towards superconductor research

London

THIS week will see a response from the UK Science and Engineering Research Council (SERC) to proposals that more resources should be focused on high-temperature superconductor research. In particular, pressure is being applied both to the SERC and to the Department of Trade and Industry by the government's chief scientist, John Fairclough, to use this research topic as a guinea-pig for multi-disciplinary research centres jointly funded by research councils and by industry. Some scientists are concerned at the speed of such a development and the possible implications for university research in general. An invitation to universities to bid for such a centre is expected from SERC by the end of this month, with bids to be submitted by the end of September.

There is speculation that such a hasty timetable is to allow applications for extra funds to be made to the Treasury at the next opportunity in November. There is expected to be strong resistance from SERC to the idea that money for such a scheme should be diverted from other programmes — indeed, the SERC is also considering an initiative for special funding of this topic to the tune of £3 million from its science and engineering boards. It is not yet clear whether government or industry will provide extra funds, although several companies are already putting resources into their own research on high-temperature superconductors.

The research centre or centres would combine the expertise of physicists, chemists, materials scientists and engineers. In this respect the centre is seen by the Cabinet Office as an ideal precursor to future centres devoted to other industrially relevant topics such as surface science and catalysis, semiconductor science or materials processing. They could be hosted by one university but involve collaboration with others. How new staff would be funded, and under what terms, would also depend on the University Grants Committee, which traditionally provides for permanent university staff and facilities.

The idea of such centres is politically sensitive in view of several proposals under discussion for a concentration of university research effort. Earlier this year the Oxburgh report on Earth science research recommended a three-tier system of departments in which major departmental research centres would be established while other departments either collaborated with such departments or were restricted to teaching only.

A yet-to-be published report from the Advisory Board for the Research

Councils is expected to propose that entire universities be ordered within such a hierarchy, while the University Grants Committee is committed to an Oxburgh-type review of university research in physics and chemistry. Thus, although new centres for superconductor research are unlikely to wait upon such proposals, they and their successors (which could follow within a year or so) may become part of a general move to concentrate and even

reduce resources. "This would be wrong-headed", commented one superconductor researcher. "Although there are advantages to be gained from a coordinated effort, it is also essential to maintain a diffuse research base, particularly given the ease with which these materials can be produced."

Universities such as Birmingham and Cambridge have already demonstrated strong interdepartmental collaborations in superconductor research and, along with Imperial College London, Oxford, Manchester, Southampton and Warwick, are gearing up to exploit the likely opportunities.

Philip Campbell

Universities joining the sprint for superconductors

Washington

ONLY eight months have passed since the discovery of high-temperature superconductors but already the first university centres dedicated to their study are appearing in the United States. The biggest financial backing appears to have been won by the State University of New York at Buffalo with a \$5 million grant from the New York state government agreed last week. But others are not far behind and may have bigger long-term plans.

The University of Houston, capitalizing on the leap in superconducting temperatures scored in the laboratory of Paul Chu, was first off the mark. Four million dollars has been found from a mixture of state, foundation and private sources for the first year of operation of a new centre, officially designated the "Texas Centre for Superconductivity" and directed by Chu.

Next, the university will go in search of federal and industrial funds. The aim is to find \$9 million a year to finance a wide range of research on the new superconductors. Serendipity played a big part in making it possible for the university to move forward quickly, according to Roy Weinstein, dean of natural sciences at Houston. By chance the university had research groups active in magnet design and vacuum-beam epitaxy alongside a strong chemical engineering department at the time Chu made his breakthrough. Some 35 professors will be involved in the new centre, each with a couple of graduate students, a postdoctoral researcher and a technician. Industry is willing to cooperate, says Weinstein.

New York State's ambitions at Buffalo are somewhat different. The new Institute for Research on Superconductivity is to focus on fostering the growth of the superconductor industry in the state of New York. Sensors, such as SQUID magnetometers, and magnets for nuclear magnetic resonance imaging will be key areas of

investigation in association with a consortium of some 20 electronics, cryogenic and chemical companies. A new building will be set up and some 150 researchers will be involved. The state backing they will receive will help to provide 'leverage' to obtain federal funds according to the new institute's director, David Shaw, professor of electrical engineering and computer engineering at Buffalo. But plenty of other universities are also likely to be chasing federal funds.

A group at Massachusetts Institute of Technology (MIT), headed by Simon Foner, is in the early stages of a proposal for a new centre that would involve many of the universities and manufacturers of the north-eastern United States. A large-scale regional centre would carry out research on both conventional superconductors and the new ceramic superconductors and, says Foner, may be "an ideal thing for one of the new science and technology centres" the National Science Foundation is seeking to establish.

Another approach has been taken by Lehigh University in Pennsylvania where an industry-university consortium goes into action on 1 August. At least eight and probably a dozen companies will each pay \$25,000 a year to support generic research which might eventually help industry. Extra funds will come from the state.

Not all researchers are thinking of new centres for their universities. Some are casting a look back at the early days of biotechnology and asking if it is not too soon to start their own companies. MIT is the first of the institutions actually to have given birth to a new business. American Research and Development, an old established venture capital firm, has set up American Superconducting around a licence for a method of making flexible superconducting wires that was developed by MIT researchers Gregory Yurek and John Vandersande.

Alun Anderson

Thatcher's science policy takes a step towards 'privatization'

London

HAVING already disposed of several nationalized industries by selling them to the general public, the British government has now embarked on an unusual scheme for gathering advice on its conduct of science policy from a newly created private organization. Last week, the Advisory Council for Applied Research and Development (ACARD), announced the formation of an independent "Centre for Exploitable Areas of Science and Technology", which will give advice on the direction of British technology to the government and, occasionally, to the general public. The new centre, described as an independent organization, will be supported by £5 million during its first five years. Of the total sum, the government has contributed £1 million, the remainder having been raised by subscription from 15 private industrial companies. For the time being, government officials will not say which they are.

Sir Francis Tombs, ACARD's chairman who is also the chairman of the newly privatized aeroengine company Rolls-Royce, is understood to have been attracted by the idea of an independent source of research and development strategy since the publication last year of ACARD's report *Exploitable Areas of Science*, in which the founding of an independent centre was a central recommendation. The scheme has also been supported recently by the Cabinet Office.

How the centre will function is nevertheless far from clear even at this stage. A statement of congratulation from the prime minister, Mrs Margaret Thatcher, says that it will monitor both research and development and changes in world markets, judge what holds promise of commercial exploitation, "find better ways of linking with scientists, manufacturers and commerce" and encourage industry and investors to keep scientists and researchers in touch with market needs.

Some officials believe that the essence of the centre's role will be to identify the directions in which the pattern of British civil research and development can most profitably be guided, using the experience of industrial and commercial people to distinguish between proposals for catching up with developments already successful elsewhere and those that could be distinctively successful. How such opinions would be communicated to the world at large remains unclear, as does the composition of the centre and its governing council, though it has been announced that Sir Robin Nicholson, the science adviser in the Cabinet Office until his move to the glass-maker Pilkington more

than a year ago, is the chairman of the steering committee.

An official says that the governing council will consist of representatives of those who have contributed to the initial costs. Asked whether academic researchers would be represented, he said that this would probably take the form of representation from the Advisory Board for the Research Councils.

Ironically, the formation of the new centre comes on the heels of the decision

last month that the Technical Change Centre, another independent centre, should be closed down due to lack of continuing support from the public research councils. The Technical Change Centre was funded eight years ago by the Leverhulme Foundation, the Science and Engineering Research Council and the Economic and Social Research Council (ESRC). Earlier this year, ESRC decided that its block grant to the centre would have to be replaced by project grants in which the centre would compete with other applicants. A government official says that the closing of one centre and the opening of the other was a "coincidence".

John Maddox

British university research has been sold "too cheap"

London

BRITISH universities have been underpricing their research effort for decades and must establish a more efficient method of estimating the true cost and charging their clients accordingly. These are the conclusions of the Committee of Vice-Chancellors and Principals which has just published two discussion papers to help create a more professional framework for the financial management of British universities.

Poor management and bad costing within the universities, claims the committee, has meant that research carried out for industrial clients and government

agencies is being subsidized from the basic grant, which is intended for pure research. The committee further calls on government to set an example by paying the "full economic cost" of the work being done. A working group of the committee which studied the problem was chaired by Professor Harry Hanham, vice-chancellor of the University of Lancaster. He says that "universities must continue to expand their income from private sources for research, short course and other projects, but it is essential that in going after this business they do not inadvertently subsidize it from their mainstream basic research and teaching budgets. If they do, the pursuit of cash today could mean intellectual and financial bankruptcy tomorrow."

The committee's views are meant to reflect the more austere climate that now exists in British universities and to endorse the encouragement given by government to increase industrial sponsorship. The vice-chancellors conclude that failure to maintain and increase the level of research funding from resources other than the University Grants Committee (UGC) and the research councils is likely to lead to a fall in the overall level of research activity and the international standing of UK universities as research institutions. The new framework for effectively costing and controlling the finances of research should include a formula for adequately rewarding the universities through royalties when their work is commercially exploited.

The universities are invited to submit their views on the proposals by 1 November. The government plans to replace the UGC with a smaller body with more industrial representation, called the Universities Funding Council (UFC), which will distribute funds to the universities "under new contract arrangements". Efficient cost control will be crucial in securing such contracts. Bill Johnstone

Park going west

London

BRITAIN'S biggest science research park has been launched in the Bristol area, and its promoters say they hope its US-style design will "help stem the brain drain" of talented scientists. The £250 million project, covering 500 acres at Emersons Green, Kingswood, is billed as "a one-stop shop for research and development, providing management, venture capital and marketing support". It will be developed in association with the universities of Bristol and Bath, and Bristol Polytechnic.

"Scientists will have the opportunity to set up their own businesses, working directly with commerce and getting the rewards they deserve," says promoter Mike Saunders. The first scientists should be working on the site within two years, he says, in an environment planned to include leisure and shopping facilities, residential units, and a hotel and conference centre.

Promoters are hoping to attract investment from high technology firms in the United States and Japan, and Mr Saunders says several international companies have expressed interest. Kathy Johnston

Japanese eye human genome

Tokyo

INTEREST in sequencing the human genome is taking off in Japan. Two international meetings have been held in the past week to discuss the issue, and the Science and Technology Agency (STA) will apply for funds to try and get the project off the ground next year.

For several years, Professor Akiyoshi Wada of Tokyo University has led a government-industry project funded by the agency to develop automated equipment for reading long DNA sequences, such as the human genome. Last week, he chaired an international workshop at Okayama, sponsored by the Hayashibara Foundation. The workshop was followed up by a one-day symposium in Tokyo on "a strategy for mapping and sequencing of the human genome" attended by several of the workshop speakers and sponsored by the STA and several Japanese companies.

An impressive array of technology was lined up by the Japanese, including a robot system developed by Seiko Instruments for DNA-sequencing based on the Sanger method, Fuji Photo Film's continuous process for manufacturing ultrathin (0.2 mm) sheets of precast gels and a laser system developed by Hitachi for reading DNA sequences on gels marked using fluorescence. Fuji Film are already marketing their gel sheets in Japan and the only major complaint from customers, says Tsutomu Hamaoka of Fuji, is the price (¥2,500, about £17 per sheet).

But the rest of the world has not been left behind. After the Hitachi speaker had finished, Lennart Philipson of the European Molecular Biology Laboratory (EMBL) pointed out that EMBL has already patented a similar laser-based system to that of Hitachi and it will be marketed shortly.

Despite such competition, many of the speakers called for international collaboration. And, although differences of opinion exist over how and when the human genome project should start, there was general agreement that the project should be a joint effort by the United States, Japan and Europe.

The STA will apply next month for a few hundred million yen (a few million dollars) to put together a fully automated sequencing system using machines developed by Seiko, Fuji and Hitachi and to develop techniques for handling and dissecting chromosomes. Further discussions will also be held with the US Department of Energy on the possibility of US-Japan collaboration. But to mount a truly international project will require much more funding.

David Swinbanks

Titanium content brings Vinland map back into play

London

THE Vinland map, which came to light in the early 1960s, and was believed for some years to be the earliest cartographic evidence of the existence of North America, was pronounced a forgery in 1974, when analysis of microparticles removed from its surface led McCrone Associates of Chicago to conclude that the ink used was based on titanium oxide, and hence of twentieth century origin. But now work by a team at the Crocker Nuclear Laboratory of the University of California (published in a recent issue of *Analytical Chemistry*) has called into question the McCrone results, and once more opens up the question of the map's authenticity.

Using a proton milliprobe, collimated to 1.0 mm × 0.5 mm (because the width of the lines on the map is of the order of 0.5 mm), Cahill and his team carried out 159 analyses of the parchment and ink of the map. According to the McCrone report, the alleged forger had used a titanium-based pigment to simulate the stain left by fifteenth century inks that have had their black layer flaked away by time. But according to the Cahill team, about one-third of all such lines on the map showed no titanium whatsoever above the minimum detectable level of 0.2 ng cm⁻².

Cahill stresses that the McCrone conclusions are based on the X-ray diffraction pattern and four electron diffraction patterns plus qualitative observations of optical behaviour. His own team, on the other hand, not only made a considerably more extensive study of the map but also

of undoubtedly authentic fifteenth and sixteenth century manuscripts, and specimens prepared using modern ink applied to a seventeenth century parchment of similar elemental composition to that of the map.

Although, in effect, refuting the McCrone report, Cahill and his team stress that they do not claim that the map is therefore authentic. Further research is needed, they say, although the obvious test, carbon-14, might be inconclusive, as it would give a dating for the parchment, not the drawing.

What is needed, too, is further cartographic and historical research on the content of the map — a sensitive subject, incidentally, in the United States, where citizens of Italian descent feel that the evidence of the Icelandic sagas for a land-fall in "Vinland" (North America) somehow detracts from Columbus's achievement. The real problem of the map for many Scandinavian scholars is Greenland, which seems simply "too good". The Viking settlers established only two small colonies in the south but the map gives, in effect, only a slightly distorted version of the true outline of Greenland, which if not a very lucky guess would seem to call for circumnavigation and sophisticated high-latitude survey instruments. Furthermore, the consensus of Scandinavian geographers of the alleged period of the map, and indeed considerably later, was that Greenland was not an island but part of the European landmass, curving eastwards until it linked up with northern Russia.

Vera Rich



Part of the Vinland map, showing North America (left), presumably as deduced from the Viking sagas. Greenland is mapped too well for the taste of Scandinavian scholars.

Fraud and other matters?

SIR—While reading the article (*Nature* **325**, 207; 1987) by Stewart and Feder, I found myself increasingly disturbed by issues peripheral to that of direct fraud.

Stewart and Feder seem to have forgotten that the 18 "full-length" research papers were supposedly peer-reviewed. Ought not the reviewers to have questioned the "errors" and "lapses" identified by Stewart and Feder (for example, the ages of parents and siblings in ref. 87)? The peer-review system is designed to prevent such things from occurring. It appears to be failing miserably, which I find more disturbing than any of the allegations made by Stewart and Feder. Perhaps the scientific community would do better if it turned its attention away from sensationalized individual instances and addressed this commonplace but probably more damaging malady.

Equally alarming is the assertion by Braunwald in his response (*Nature*, **325**, 215; 1987) that the republication of abstract 70 as abs. 78 is legitimate on the grounds that abs. 70 had been rejected. I have personally undertaken a survey of medical and dental school deans, all of whom unanimously agreed that, to their knowledge, rejected abstracts are not published by *Clinical Research*, *Circulation* or any other journal. Braunwald put himself in the position of answering an allegation of dishonesty with another. How was abs. 70 both published and rejected by the American Federation for Clinical Research? As one dean asked, is the inclusion of rejected abstracts, manuscripts and funding proposals in one's *vitae* a standard practice at Harvard? If such an institution condones such practices, how can one of its junior faculty members (Darsee) be criticized for bending other conventions?

MICHAEL J. GLADE

Northwestern University,
Medical School,
Department of Pharmacology,
303 E. Chicago Ave.,
Chicago, Illinois 60611, USA

BRAUNWALD REPLIES — Abstract no. 70, which Michael Glade refers to, was submitted for presentation to the American Federation for Clinical Research. *Clinical Research*, the journal of this society, has the unusual policy of publishing all submitted abstracts, regardless of whether or not they are accepted for oral presentation. Indeed, only a fraction of submitted abstracts are accepted for oral presentation. This particular abstract was not accepted for oral presentation. The abstract was slightly revised and then submitted for consideration for presentation before the Scientific Sessions of the American Heart Association and was accepted for oral presentation and pub-

lished in *Circulation* (Abstract no. 78). This sequence is not considered inappropriate by either society.

I am not aware that I or anyone at Harvard has included rejected manuscripts in either *vitae* or funding proposals and am not sure what Glade was referring to.

EUGENE BRAUNWALD

Brigham and Women's Hospital,
75 Francis Street, Boston,
Massachusetts 02115, USA

Primary misnomer

SIR—I should like to call attention to the inappropriate usage of 'primary sequence' as a term to describe the first level of protein structure. By all conventions this phraseology is a misnomer; certainly, 'logical' extensions such as 'secondary sequence' and 'tertiary sequence' are fictitious. Rather, this element of protein structure should be referred to as 'amino-acid sequence', 'primary structure' or 'covalent structure' (see, for example, *J. biol. Chem.* **251**, 11–12; 1976).

My concern stems from the alarming increase in frequency with which this misnomer is being used at major scientific meetings and in many internationally recognized publications. For example, a computer-assisted search of the literature for the past three years shows that 'primary sequence(s)' has appeared in journal titles at least 18 times and in abstracts at least 140 times; the journals 'represented' frequently include *Nature*, *Science*, *Proceedings of the National Academy of Sciences of the USA*, *Journal of Biological Chemistry*, *Nucleic Acids Research*, *EMBO Journal*, *Journal of Virology*, *Biochemistry*, *Journal of Molecular Biology* and *FEBS Letters*.

SHAUN D. BLACK

Division of Medicinal Chemistry &
Pharmacognosy,
College of Pharmacy,
Ohio State University,
Columbus, Ohio 43210, USA

Dating the Shroud

SIR—My first reaction to the letter from Denis Dutton (*Nature* **327**, 10; 1987) was one of satisfaction that concern should be expressed about the protocols for the tests to carbon-date the Shroud of Turin.

However, on reflection I realized that his remarks were a gross insult to the 22 experts who met in Turin between 29 September and 1 October 1986 to work out the procedure to be used in carbon-dating the Shroud. The remarks also cast a shadow on the integrity of the guarantor from the British Museum and the seven laboratories taking part in the tests

(*Nature* **323**, 486; 1986). It implies, too, that the 40 scientists involved in the STURP investigations of 1987 would stand by and allow their credibility to be sacrificed by some sleight-of-hand trick with the samples.

Then I recalled that the US Committee for the Scientific Investigation of Claims of the Paranormal has a long history of hostility to the claims made for the Shroud of Turin. Unfortunately, the committee can hardly be regarded as having a reasonable and open-minded approach to the subject.

The problem facing Dutton and the US committee is that they fail to realize that proving the shroud to be a mediaeval forgery no more disproves the existence and claims of Jesus Christ than does the discovery, for instance, that a particular lock of hair could not have come from Napoleon would disprove that he existed.

However, if the Shroud is shown to be a mediaeval forgery, science will still be left with the question of how a mediaeval artist produced a three-dimensional photographic image of a body in an advanced state of *rigor mortis* complete with anatomical details that can be recognized by physicians and forensic pathologists.

Thus whatever the results of the carbon-dating of the shroud, there will be much work for dedicated scientists, as distinct from stage magicians, to carry out to resolve the mystery of the shroud. Whatever the result, I feel that the "deep religious passions" will be displayed by the Denis Dutton school of thought and not by the scientific community, whose aim is to establish the truth.

P.R. SMITH

CSIRO Division of Mineral Chemistry,
PO Box 124,
Port Melbourne,
Victoria 3207, Australia

How no green cow

SIR—Ralph A. Lewin's question "Why are cows not green?" (*Nature* **326**, 743–74; 1987) can be generalized: why are there no green mammals? Green would be the optimal colour for camouflage in grassland and forest environments, and there are plenty of green birds, reptiles and amphibians, not to mention invertebrates. For that matter, few mammals have bright colours of any kind. Can we have inherited from drab nocturnal ancestors deficiencies in our system of pigmentation so deep-seated that tens of millions of years of evolution have failed to correct them? Perhaps coloured clothing was the first conspicuous triumph of cultural developments.

PHILIP J. STEWART

University of Oxford,
The Pauling Human Sciences Centre,
58 Banbury Road,
Oxford OX2 6QS, UK

But how thin is a thin wire?

Theory and measurement conspire to show that there are quantum limitations of the regularity with which thin wires conduct electricity.

ATOMIC engineering is not necessarily the craft of designing nuclear reactors. Not so long ago, people were talking optimistically of building microelectronic devices on the scale of interatomic distances, and for obvious reasons: the smaller the devices, the faster will be the arrays into which they are incorporated. But it has always been on the cards that the ambitions of the device constructors would be frustrated by basic difficulties, quantum mechanics for example.

Two neat experiments show that indeed to be the case. It may not be a coincidence that the two groups are from IBM's Yorktown Heights Laboratory (Benoit, A., Umbach, C.P., Laibowitz, R.B. & Webb, R.A. *Phys. Rev. Lett.* **58**, 2343; 1987) and from AT&T Bell Laboratories (Skocpol, W.J., Mankiewicz, P.M., Howard, R.E., Jackel, L.D. & Tennant, D.M. *Phys. Rev. Lett.* **58**, 2347; 1987). The Bell team worked with A.D. Stone of Yale University, one of the pioneers in this field.

The question is what happens to an electric current flowing in a wire so thin that quantum effects become important? It is tempting to think of a current as a plane electron wave propagating through the wire, but that cannot be the case when the wire has resistance because electrons are occasionally scattered (losing energy, which is why the wire gets hot). So why not look for quantum interference between scattered electrons? This expectation has been confirmed in the past two years by placing thin wires carrying a constant current in a magnetic field and showing that the measured voltage drop, or measured resistance of the wire, varies with the magnetic field.

But when is a wire properly called "thin"? On the view that a current in a superconducting wire is a plane electron wave travelling along, and that a current in a normal wire is a plane made incoherent by random scattering, a wire can be called thin when its width is less than the length over which the coherence of an electron wave is destroyed by random incoherent scattering. That seems to be the starting point for the measurements now described. They work with conductors with lateral dimensions of a micron or a tenth as much, and even so must use liquid-helium temperatures for the quantum effects to become apparent.

Neither group of authors ducks the interesting question of just what is meant by the measurement of a voltage along the

length of such a wire. Obviously, it is not possible to attach voltage probes mechanically to wires made, in one case, by evaporating either gold or antimony by a lithographic process in which a scanning electron microscope was used to etch the pattern. Instead, the probes must be integral with the conductor, which graphically shows that, if the measuring instrument is physically a part of the experiment, measurement can hardly avoid interfering with the phenomenon observed. A voltmeter is a device for measuring the difference of the chemical potential of electrons within the two conducting attachments and must be accessible to electrons and thus contribute to the loss of coherence.

Quite what resistance means is another proper question. Plainly it is not safe to talk about the resistance per unit length of the conductor, but only about the apparent resistance between two specified probes, calculated from Ohm's Law as the ratio of the measured current to the known voltage, thus allowing that the result may be determined by the shapes and other properties of the probes. Both sets of experiments use conductors equipped with at least four built-in probes which can be joined together in six different pairs for voltage measurements. The trick is to span the expected incoherence length with the range of distance separating the possible pairs of probes.

The new result, at which both groups arrive, concerns what happens with the thinnest wires, but the Bell group also neatly confirms what is expected to happen in other circumstances. Given the imponderables, there is no point in simply measuring the resistance between all possible pairs of points; because of quantum interference that will affect the different probes differently, it will not be possible accurately to relate the apparent resistance between a pair of probes with the length of the conductor between them, which is where the random disturbances caused by external magnetic fields are useful: the magnitude of the fluctuation of the voltage between two specified probes is a measure of the extent to which quantum interference affects that part of the system, and is thus a measure of the extent to which a voltage difference is affected by quantum interference even in the absence of a field.

First, there is a simple and almost classical argument. Suppose you have a length

of thin wire and place voltage probes supposed not to interfere with the phenomenon observed at equal lengths along it. If there is quantum interference affecting the voltage differences between successive pairs of probes, it will at least be possible to relate the fluctuation along the total length to the fluctuations across the separate pieces by exactly the argument used to show that the standard error of N measurements affected by a source of random error is inversely proportional to the square-root of N . Because the voltage will on the average be proportional to the length, the proportional departure of the voltage from expectation will be inversely proportional to the length of the wire. In other words, as common sense suggests, the voltage drop along a long thin wire will be relatively unaffected by the effects of quantum interference at the voltage probes.

The new development is that this appears not to be the case when the distance between the voltage probes is smaller than the incoherence length. Then, remarkably, it seems that the variation of the apparent resistance is constant, and independent of the distance between the probes and thus of the expected value of the resistance. Both groups have shown this by measurement. The IBM group has also inferred that this must be the case by means of a neat use of statistical arguments and an inference about the symmetry of various aspects of the fluctuations with respect to reversals of the magnetic field derived from Onsager's reciprocity relations.

This result is not encouraging for the designers of microelectronic devices. When they embark on making microcircuits whose elements are shorter than the incoherence length, they will discover that mere geometrical length is no longer a good guide to electrical properties. Moreover, it will then emerge that the relative variation of the error due to quantum interference becomes a larger proportion of the total resistance as the length of the circuit element shrinks. It is easy to imagine what disastrous effects this would have on the performance of a microcircuit with, say, a million conducting lengths of thin wire, each with a resistance varying randomly and unavoidably. The only good news is that on the face of things these gloomy speculations should not apply to superconductors.

John Maddox

Astronomy

A means to sharper images

E. Keith Hege

INTRICATE details of the astronomical systems observed with Earth-bound telescopes are destroyed by atmospheric distortion. But advances in our understanding of the Earth's atmosphere, together with recent image-reconstruction techniques, promise to make available distortion-free images commensurate with the light-gathering capabilities of modern apparatus. On page 229 of this issue, L.A. Thompson and C.S. Gardner describe an artificial star, produced by a laser beam scattered in the upper atmosphere, that may enable astronomers to unscramble atmospheric effects in distorted images. This technique could have applications in high-energy-beam defence systems.

Until recently, the giant telescopes of modern astronomy, used as cameras that record the images of distant objects, could see no more sharply than Isaac Newton's telescope. The large instruments are used to gather more light, and thus to look at greater distances, but their intrinsic ability to produce sharp images is seriously impeded by the turbulence of the atmosphere, through which all ground-based instruments must view the sky. This turbulence limits the sharpness of a directly-recorded image from any large telescope to that of a small telescope with an aperture of only 10–15-cm diameter, except during the very rare moments of exceptionally low turbulence, which astronomers call 'good seeing'. Even in those rare moments the resolution is insufficient to reveal details on the surface of the nearest stars, other than the Sun. The full, diffraction-limited resolution of the largest ground-based telescopes can be exploited, however, by using fast computers that execute complex image-reconstruction algorithms. For this, short-exposure images are required, obtained with highly efficient detectors capable of recording individual photons.

These image-reconstruction methods are limited to objects of sufficient brightness that several correlated photons can be recorded in the time during which the atmospherically distorted image does not change by more than the telescope's limit of resolution — typically 5–50 ms. Correlated photons are those that arrive within the 10–15-cm subaperture of the telescope's pupil that determines the sharpness produced in the conventional camera mode. These requirements limit these methods to sources brighter than visual-magnitude 17, independent of the size of the telescope, corresponding only to the closest galaxies with active nuclei

and the brightest quasi-stellar objects.

Fortunately, optical physics enables us to unscramble the atmospheric distortions if the image field of view contains an unresolvable (point) source, but very few objects of interest are matched with such bright point sources. The field of view to which any of these methods can be applied, the so-called isoplanatic field of view, is very limited — typically about 2 arc seconds. For the cases in which a sufficiently bright point is present, two methods are available to remove or reduce the atmospheric blurring. Adaptive-optics systems can measure and correct for the atmospheric distortion; deconvolution methods, related to holography, can correct a sequence of simultaneous short-exposure images of the point source and the object of interest. If the point source is bright enough, it should be possible to correct the image of any source in the isoplanatic field of view.

The results of the experiment reported

by Thompson and Gardner are therefore of great relevance in the quest for high-resolution images. The artificial star they create by scattering a laser beam off sodium vapour in the upper atmosphere is an alternative to a genuine natural point source that can always be put in the field of view of the telescope.

It seems unlikely that it will be possible to project a 'star' sufficiently point-like to be the perfect key to deconvolution. The mesosphere is neither infinitesimally thin nor infinitely distant, and the 'star' will have a size determined by diffraction in the projector itself (which could be the telescope). Nevertheless, the experiment shows that it is possible to project a source that is sufficiently bright to be useful. The approximate key to deconvolution provided by the guide star should increase the size of the effective correlated aperture and the effective atmospheric correlation time. These two benefits, in a system using adaptive optics to provide optimum data to the image-reconstruction methods otherwise available, can extend the highest resolving powers to even fainter targets in the early Universe. □

E. Keith Hege is at the Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA.

Molecular neurobiology

Channel families in the brain

Charles F. Stevens

MANY in the molecular neurobiology business have suspected all along that channels — membrane proteins mediating transmission of information in the brain — would fall into genetically related families, but this expectation has been based on analogies to other systems, such as the rhodopsin/ β -adrenergic/muscarinic receptor family¹, rather than on direct information. Now, suddenly, it has been shown in four different laboratories that the expectation is correct and family structures are beginning to emerge. Specifically, the glycine (page 215 of this issue²) and GABA_A (page 221³) receptors share similarities in amino-acid sequence with the nicotinic acetylcholine receptor channel, and a mammalian muscle calcium channel⁴ and a *Drosophila* (fruitfly) potassium channel⁵ are similar to the sodium channel.

Up until now, the taxonomy of channels has been based on their functional properties. Two main groups of channels have been identified: ligand-gated channels, proteins located in postsynaptic membranes that bind specific neurotransmitters and open a pore to function in synaptic transmission; and voltage-gated channels, integral membrane proteins designed to sense the transmembrane electrical field and open ion-specific pores to produce

nerve impulses and other forms of electrical excitability. Each of these two groups has perhaps dozens of members. In the past, ligand-gated channels were classified pharmacologically (according to which substances they bind), whereas the voltage-gated channels were distinguished by the ionic selectivity of their pores and by properties related to the manner in which they open and close. The new data mean that channels can be classified according to their structure.

The first ligand-gated channel to have its structure determined was the nicotinic acetylcholine receptor (AChR) (see the letter to *Nature* by Noda *et al.*⁶). Various AChRs, both from muscle and from brain, are known, and these molecules themselves form a family tied together by strong similarities. Their important characteristics are that each receptor is constructed from multiple distinct subunits (four in muscle but perhaps only two in brain^{7,8}) that are closely related genetically; and each subunit is structurally similar to others in that they all have a very similar pattern of alternating hydrophilic and hydrophobic regions (*a* in the figure). In particular, there are four domains, each with around 20 consecutive hydrophobic amino acids, in all of the known AChR

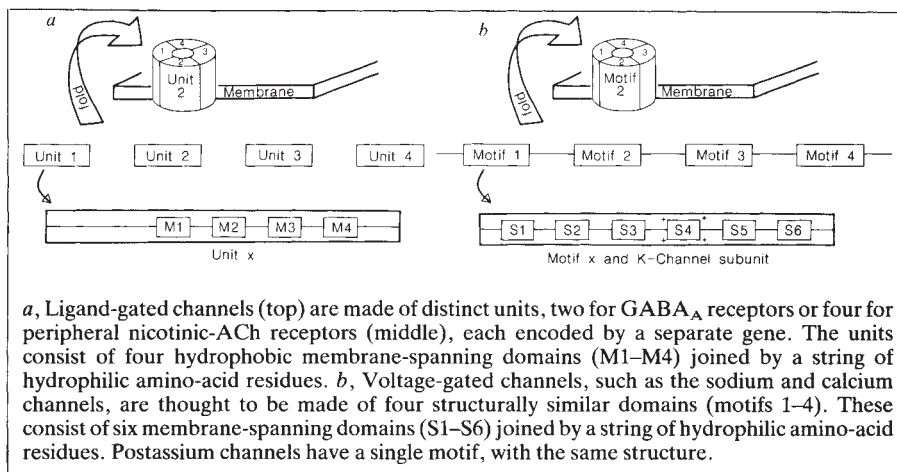
subunits from all species, and these domains are thought to constitute membrane-spanning helices.

Three peptides from ligand-gated channels of the central nervous system have now been sequenced, the two subunits of a GABA_A (γ -aminobutyric acid) receptor channel by Eric Barnard's group (see page 221³) and one from a glycine receptor channel by Heinrich Betz's group (see page 215⁴). Both these chloride channels mediate inhibitory synaptic transmission, whereas the AChR sodium channel mediates excitatory transmission. What is striking is that all three subunits are very similar to each other, with about 50 per cent homology (that is, about half of their amino-acid residues either are identical or are substitutions frequently made in protein evolution), and all about 25 per cent homologous to those of the AChR. Further, all have the same pattern of four putative transmembrane domains, and the GABA receptor channel at least is like the AChR in that its subunits are closely related; presumably the same will be true for the glycine receptor channel when the structure of its other subunit is elucidated.

Unlike the polymeric AChR, the sodium channel — the first voltage-gated channel to have its sequence determined — can function with a single polypeptide chain about four times as large as any of the AChR subunits. This polypeptide chain contains four very similar motifs, each perhaps analogous to a single subunit of a multimeric protein (*b* in the figure). In turn, each of these motifs has the same characteristic pattern of alternating hydrophobic and hydrophilic regions that are interpreted as five membrane-spanning helices. In addition, and most interestingly, another region found in each motif has the general structure (Arg-X-X)_n, where X is a hydrophobic residue. This structure is long enough to make a membrane-spanning helix. This region, called S4, is thought to be the voltage sensor that couples transmembrane electric field to the conformational change in the protein that opens and closes the pore of the channel.

Now two more proteins, thought to be a muscle calcium channel and a *Drosophila* potassium channel, have had their sequences determined by Numa's group⁴ and by the Jans' group⁵, respectively. The calcium channel is very like the sodium channel: it is a long polypeptide with four motifs, each of which has the same pattern of putative membrane-spanning regions and each of which has the characteristic S4 regions with the repeated Arg-X-X. Overall, the sodium and calcium channels have about 60 per cent homology, clearly two rather closely related family members.

Perhaps even more interesting is the potassium channel called, for historical reasons, the A-current channel. This protein is about a third as large as the sodium and calcium channels and seems to be just



a, Ligand-gated channels (top) are made of distinct units, two for GABA_A receptors or four for peripheral nicotinic-ACh receptors (middle), each encoded by a separate gene. The units consist of four hydrophobic membrane-spanning domains (M1–M4) joined by a string of hydrophilic amino-acid residues. *b*, Voltage-gated channels, such as the sodium and calcium channels, are thought to be made of four structurally similar domains (motifs 1–4). These consist of six membrane-spanning domains (S1–S6) joined by a string of hydrophilic amino-acid residues. Potassium channels have a single motif, with the same structure.

a single copy of the motif that is represented fourfold in these other channels. Only a single part of the A-current channel sequence is similar to the sodium/calcium channel: the hundred or so amino acids surrounding an S4-like region are about 50 per cent identical. Even more striking than the conservation of this presumed voltage sensor is the hydrophobicity pattern. Although most of the amino-acid sequence is unrelated to that of the sodium/calcium channel, the pattern of alternating hydrophobic and hydrophilic stretches is amazingly similar. As the potassium channel is thought to be phylogenetically old⁹, it may be an ancestor of other more modern voltage-gated channels that have evolved by gene duplication into their present forms. Some support for this view comes from the recent observations of Salkoff *et al.*¹⁰ who show that the *Drosophila* sodium channel has the same four-motif structure as, and

has a high degree of homology to, the vertebrate sodium channel.

The race is now on. With the identification of conserved regions that define family membership, people in many laboratories will start searching with synthetic oligonucleotide probes for yet other family members, and we can anticipate that the next few years will finally provide the data for a rational taxonomy of channels in the brain. □

1. Hall, Z.W. *Trends Neurosci.* **10**, 99–101 (1987).
2. Grenningloh, G. *et al. Nature* **328**, 215–220 (1987).
3. Schoffield, P.R. *et al. Nature* **328**, 221–227 (1987).
4. Tanabe, T. *et al. Nature* (in the press).
5. Tempel, B.L. *et al. Science* (in the press).
6. Noda, M. *et al. Nature* **302**, 528–532 (1983).
7. Whiting, P.J. *et al. Nature* **327**, 515–518 (1987).
8. Boulter, J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
9. Hille, B. *Ionic Channels of Excitable Membranes* (Sinuar, Sunderland, Massachusetts, 1984).
10. Salkoff, L. *et al. Science* (in the press).

Charles F. Stevens is in the Department of Molecular Neurobiology, Yale University Medical School, New Haven, Connecticut 06510, USA.

AIDS

Infectious, genetic or both?

Jared M. Diamond

SOME of the people exposed to HIV (human immunodeficiency virus) by unprotected sexual intercourse or by blood transfusion fail to become infected¹. Among those infected, there is much variation in progression to full-blown AIDS (acquired immune deficiency syndrome). Although differences in conditions of exposure undoubtedly contribute heavily to those individual differences in susceptibility, a recent study by Eales *et al.*² now suggests that a genetic factor contributes as well. AIDS, the paradigmatic infectious disease of the late twentieth century, thus joins the growing number of conditions that blur the traditional distinction between infectious and genetic diseases.

The suggested genetic factor in this case is a protein called Gc (group-specific complement) or vitamin D-binding factor, which has three common, widespread

alleles termed 1F, 1S and 2. Among British men the allele *Gc1F* is reported to be positively associated, and *Gc1S* and *Gc2* negatively associated, with three factors: the risk of infection following exposure to HIV; the severity of clinical symptoms; and the progression to AIDS. No AIDS patient in the studied sample was *Gc2* homozygous, but no subject seronegative for HIV (despite known contact with AIDS patients) was *Gc1F* homozygous. Thus, *Gc2* apparently protects against AIDS, whereas *Gc1F* predisposes to it. It remains unknown whether the allele differences relevant to HIV infection involve differing amino-acid sequences or post-translational changes in content of sialic acid, which could be involved in HIV binding^{2–5}.

The study by Eales *et al.* is stimulating efforts by others⁶ to test the claimed

association between Gc and AIDS and to search for other possible genetic associations. The frequencies of Gc alleles show marked geographical variation, which could be clinically significant if the association between AIDS and Gc is confirmed. Thus, sub-Saharan mainland Africa has one of the highest reported frequencies of *Gc1F* (0.58–0.88) and one of the lowest frequencies of *Gc2* (0.02–0.09)^{3,7}. Perhaps this high African frequency of *Gc1F* is among the factors that contributed to the spread of AIDS there.

To appreciate the broader significance of these observations, recall that genetic diseases and infectious diseases used to be the province of separate specialists, departments and textbooks. But this distinction has been crumbling with each new discovery of genetic susceptibility underlying infectious diseases. For instance, tuberculosis, the paradigmatic infectious disease of the nineteenth century, although undoubtedly caused by infection with the tubercle bacillus, is also co-determined by genetic resistance factors that modulate the immune response and gastric acidity^{8,9}. Discoveries of genetic susceptibilities have similarly been eroding the line between genetic diseases and environmentally triggered diseases. Although smoking is a paradigmatic environmental cause of disease, for example, genetic factors contribute to the great individual variation in susceptibility to smoking¹⁰.

In retrospect, two reasons for this blurring of distinctions could have been anticipated. First, no infectious disease — not even the Black Death in mediaeval Europe, nor smallpox on first contact with American Indians — has killed all of a human population. Survival can be influenced by many factors, including nutritional and environmental as well as genetic ones. Second, whenever genetic factors thus contribute to differential mortality from infection, natural selection will cause the protective genetic factors to increase in frequency as long as the infective agent remains present. For instance, natural selection led to a high frequency of sickle-cell haemoglobin (HbS) in the African malarial belt through selection of HbS heterozygotes for their resistance to malaria. Subsequently, through selection against HbS homozygotes because of anaemia, natural selection led to a decline or virtual disappearance of HbS in New World blacks or South African blacks, respectively, whose ancestors carried the *HbS* gene into those malaria-free regions (where the gene lost its advantage)¹¹. These two types of reasons are opposite faces of the same coin: genes have a major influence on susceptibility to infectious disease, whereas infectious disease is a major selective force on gene frequencies.

In this light, it is instructive to consider how various specialists describe the

disease patterns of Yupik-speaking Alaskan Eskimos¹². Bacteriologists have measured the world's highest incidence of a severe infectious disease, invasive *Haemophilus influenzae* type b (HIB) disease, in this population. Geneticists, however, have observed the world's highest incidence of a genetic (autosomal recessive) disease, salt-losing congenital adrenal hyperplasia (CAH) caused by 21-hydroxylase deficiency. But those Eskimos carrying a genetic marker for CAH turn out to be virtually unrepresented among the HIB patients, who make up nearly half the studied population. Thus, an evolutionary biologist would conclude that severe HIB infections among Eskimos had selected for genetic resistance factors, and that CAH is a price paid by homozygotes for escaping HIB.

What does the future hold for geneticists interested in AIDS? If AIDS kills a significant fraction of the human population, it will provide a unique and grisly opportunity to document natural selection in humans. Genetic susceptibility factors underlying other severe infectious diseases were identified only after natural selection had already skewed gene frequencies by decimating susceptible people. Thus, as in the classic example of HbS and malaria, natural selection could only be inferred by hindsight. Changes in allele frequencies were not followed with time, as will now be possible with AIDS.

The frequencies of Gc alleles were

measured in samples collected well before the AIDS epidemic reached Europe and North America. Two facts^{2,7} show that AIDS has as yet killed too low a proportion of susceptible people to skew gene frequencies: the frequencies of Gc alleles in the British homosexual and heterosexual samples were the same; and the frequency of the AIDS-predisposing allele *Gc1F* is still three times higher in central Africa (where HIV infection is common) than in Europe. Will *Gc1F* frequencies plummet and *Gc2* rise, as susceptible people die and as those with genetic resistance survive? If HIV becomes so widespread that everyone somehow gets exposed, will the matter of who develops AIDS come to depend as much on differences in genetic factors as on differences in exposure, and will we then view AIDS as a genetic disease? □

1. Pinching, A.J. *Clin. Immun. Allergy* **6**, 467–488 (1986).
2. Eales, L.-J. *et al. Lancet* **i**, 999–1002 (1987).
3. Kamboh, M.I. & Ferrell, R.E. *Hum. Genet.* **72**, 281–293 (1986).
4. Whitehouse, D.B. *et al. Lancet* **i**, 1377 (1987).
5. Bowman, B.H. *Lancet* **i**, 1377–1378 (1987).
6. Thymann, M. *et al. Lancet* **i**, 1378 (1987).
7. Eales, L.-J. *et al. Lancet* **i**, 1268–1269 (1987).
8. Youmans, G.P. *Tuberculosis* (Saunders, Philadelphia, 1979).
9. Peterson, G.M. & Rotter, J.I. *Am. J. Phys. Anthropol.* **62**, 71–79 (1983).
10. Hankins, D. *et al. Am. Rev. Respir. Dis.* **125**, 119–121 (1982).
11. Beighton, P. & Botha, M.C. *S. Afr. Med. J.* **69**, 293–296 (1986).
12. Peterson, G.M. *et al. Am. J. Hum. Genet.* **36**, 177S (1984).

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Solid-state physics

Interactions of heavy electrons

M. Brian Maple

A SIGNIFICANT part of a recent meeting* was devoted to the physics of so-called heavy-electron materials, a subject that has received much attention during the past few years. The three aspects of this subject that were addressed were: the origin and nature of the heavy-electron state; the instability of the heavy-electron Fermi liquid towards the formation of superconducting and charge- or spin-ordered ground states; and the types and underlying physical mechanisms of superconductivity in heavy-electron systems. These important and timely results on the heavy-electron problem, however, were largely overshadowed by recent developments in the field of high-temperature (T_c) superconductivity which have been covered extensively in News and Views for the past few weeks. Interestingly, the mechanisms responsible for superconductivity in the high- T_c oxides and the heavy-electron materials may be similar, even though the effective mass, m^* , of the

electrons in the high- T_c superconducting copper oxide materials is less than, or of the order of, the free-electron mass m_e , whereas m^* of the heavy-electron superconductors is of the order of several hundred m_e .

Heavy-electron systems (often called heavy-fermion systems) are compounds of rare-earth (usually cerium) and actinide (usually uranium) elements in which the coefficient γ of the electronic contribution to the specific heat, $C_e = \gamma T$, is enormous, attaining values of several J mol⁻¹ K⁻² in some of the heavy-electron materials. These large values of γ , from which the large values of m^* are inferred, are accompanied by anomalies in various physical properties such as the electrical resistivity, thermoelectric power and magnetic susceptibility¹. An increasing number of people believe² that the heavy-electron Fermi-liquid state that occurs in these materials at low temperatures is a consequence of the Kondo effect — an effect in which the spins of the conduction electrons tend to screen away the spin of

* American Physical Society meeting, New York City, 16–20 March 1987.

electrons tend to screen away the spin of each magnetic ion in the lattice below a characteristic temperature called the Kondo temperature T^* . The superconducting heavy-electron compounds with the largest effective masses, CeCu_2Si_2 , UBe_{13} and UPt_3 , have attracted much attention because of the possibility that they exhibit an unconventional type of superconductivity analogous to the triplet-state (spin-parallel) superfluidity of ^3He .

One of the most interesting experimental developments during the past year is the observation of the de Haas-van Alphen (dHvA) effect, a periodic dependence of magnetic susceptibility on applied magnetic field, in heavy-electron compounds (L. Taillefer, Cambridge Univ.) because it gives direct evidence for the existence of heavy-fermion quasiparticles. The dHvA experiments on a high-purity single crystal of the heavy-electron superconductor UPt_3 reveal oscillations composed of up to eight frequency components, corresponding to cyclotron orbits in a plane normal to the a -axis. A study of their amplitudes as a function of temperature and magnetic field in the intervals 20–150 mK and 40–115 kG, respectively, yields estimates of the cyclotron masses of 25–90 m_e .

The low-temperature specific heat of various heavy-electron compounds such as CeAl_3 , CeCu_6 and UPt_3 have been measured under pressure (N.E. Phillips, Univ. California, Berkeley). In general, γ and, in turn, m^* , decrease rapidly with pressure, which is consistent with the increase of T^* with pressure generally observed in heavy-electron materials. Measurements of the Hall effect (the electrostatic field set up by a magnetic field applied to a current-carrying sample) on heavy-electron compounds such as

UBe_{13} and CeCu_6 reveal that coherent scattering of electrons has a dramatic effect on the Hall constant (T. Penney, IBM Thomas J. Watson Research Center): a high-temperature incoherent-scattering regime, a transition regime and a low-temperature coherent-scattering regime can be clearly distinguished.

Neutron-scattering measurements on the heavy-electron compounds CeCu_6 , U_2Zn_{17} and UPt_3 , which are paramagnetic, antiferromagnetic and superconducting, respectively, at low temperatures were reviewed by G. Aeppli (AT&T Bell Labs). Single crystals of CeCu_6 and UPt_3 show strong antiferromagnetic correlations, suggesting that the superconductivity exhibited by UPt_3 , which seems to be anisotropic, involves pairing of electrons in a spin-singlet state.

The Josephson current I_c between a tantalum-wire probe and an induced-surface singlet superconducting state in UBe_{13} decreases with temperature below the critical temperature for superconductivity in bulk UBe_{13} (E.L. Wolf, Polytechnic Institute of New York). This contrasts to the increase in I_c in a molybdenum reference sample, indicating that the bulk superconductivity suppresses the induced-singlet superconductivity. These tunnelling results add to the evidence for unconventional superconductivity in heavy-electron compounds, although a completely consistent picture has yet to emerge. Recent studies³ of T_c under pressure in the $\text{U}_{1-x}\text{Th}_x\text{Be}_{13}$ system indicate two distinct superconducting phases, one for $0 < x < 0.017$ and another for $0.017 > x$. Measurements of the critical magnetic field, that quenches superconductivity, as a function of temperature in samples doped with gadolinium seem to be consis-

tent with Balian-Werthamer-like triplet or singlet superconductivity for $x = 0$ and Anderson-Brinkman-Morel-like triplet superconductivity for $x = 0.029$. The thermal conductivity, ultrasonic attenuation and NMR relaxation rate all show power-law temperature dependences that suggest that the superconducting energy gap goes to zero on some parts of the Fermi surface; in conventional superconductors the energy gap is nearly constant.

The compound URu_2Si_2 is one of the most striking of the heavy-electron materials (M.S. Torikachvili, Univ. California, San Diego). Even though its m^* is only about 25 m_e , this compound is particularly interesting because it exhibits two electronic phase transitions: a charge- and/or spin-density wave transition at $T_0 = 17.5$ K; and a superconducting transition at $T_c = 1.5$ K. According to recent neutron-scattering measurements, the transition at 17.5 K is to an antiferromagnetic phase with the ion spins aligned along the tetragonal c -axis and with an ordered magnetic moment of only $0.03 \pm 0.01 \mu_B$.

Returning to the relationship between the heavy-electron superconductors and the new high- T_c superconducting oxides, both these materials may have a magnetic mechanism that is responsible for the formation of the superconducting electron pairs (Cooper pairs). Such a mechanism, involving the resonating valence bond (RVB) proposed by Pauling in 1953, was recently advanced by P.W. Anderson^{4,5} to account for the high- T_c superconductivity in the new materials. In this model, the RVB state corresponds to an insulating magnetic phase involving spin-singlet nearest-neighbour pairs of Cu^{2+} ($S = 1/2$) ions; sufficient doping to increase the Cu^{3+} concentration transforms the insulating RVB state into a metallic state in which the singlet electron pairs of the insulating state become superconducting Cooper pairs. Lines on the Fermi surface where the superconducting energy gap vanishes are also predicted that indicate, along with the antiferromagnetic correlations involving the nearest-neighbour spin singlets, similarity between the RVB and heavy-electron superconducting states. Moreover, the presence of copper in two valence states, Cu^{2+} and Cu^{3+} , is reminiscent of a class of rare-earth compounds that are related to the heavy-electron materials in which the rare-earth valence fluctuates between two integral values such as Ce^{3+} and Ce^{4+} or Yb^{2+} and Yb^{3+} . □

1. Fisk, Z., Ott, H.R., Rice, T.M. & Smith, J.L. *Nature* **320**, 124–129 (1984).
2. Lee, P.A., Rice, T.M., Serene, J.W., Sham, L.J. & Wilkins, J.W. *Comments Solid St. Phys.* **12B**, 1892 (1982).
3. Maple, M.B. *et al.* *Proc. 5th Int. Conf. Valence Fluctuations* (in the press).
4. Anderson, P.W. *Science* **235**, 1196–1198 (1987).
5. Anderson, P.W. & Abrahams, E. *Nature* **327**, 363 (1982).

M. Brian Maple is in the Department of Physics, University of California, San Diego, La Jolla, California 92093, USA.



100 years ago

Archibald's Captive Kite-Balloon.

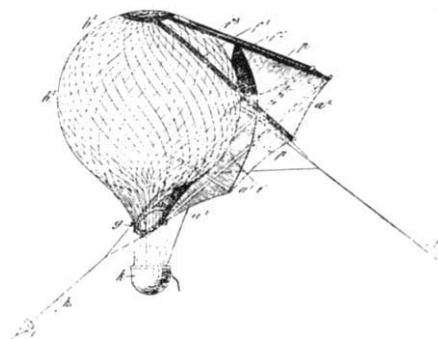
IT has always been an objection to the extensive use of captive balloons for scientific or military purposes, that a wind of moderate strength suffices not merely to depress them considerably from the vertical, but to cause them to jerk, rotate, and oscillate vertically and horizontally in such a manner as to render them either partially ineffective or totally useless.

During the recent military manoeuvres at Dover, it was stated that the captive balloon under the charge of Major Timpler was not allowed to ascent beyond the shelter of the surrounding downs, owing to the strong wind. It was thus *hors de combat* as far as the enemy was concerned, and this seems to be a common experience of military balloonists.

All the preceding defects are remedied and several positive advantages are gained by attaching a balloon to a kite in the manner

indicated in the accompanying diagram. The addition of the kite with the fastening at the side instead of the base counteracts the depression produced by the wind, and not only raises its own weight, but even in a light anticyclonic breeze elevates the whole apparatus to a higher level than that which could be attained by the balloon alone.

I arrived at the idea of uniting the two apparatuses while conducting my kite anemometrical observations in 1884, owing to my desire to prevent my kites from coming down suddenly when the wind dropped.



From *Nature* **36**, 278; 21 July 1887.

Molecular biology

Genetics of development elucidated by nematodes

Philip Anderson and Judith Kimble

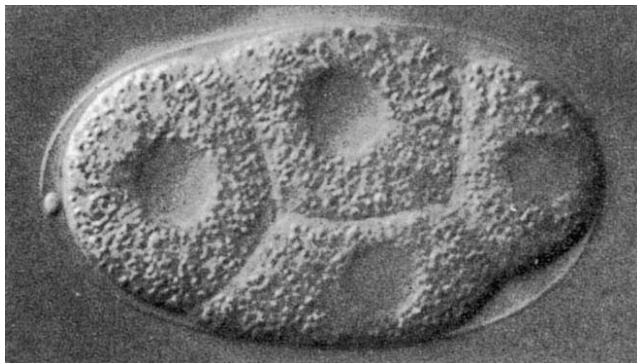
THE molecular mechanisms responsible for development of metazoan pattern and form are largely unknown. Embryos have been described and experimentally manipulated for more than a century, but only in the past few years have some of the genes and proteins that influence, and perhaps govern, development been isolated and scrutinized. These genes, cloned chiefly from the fruitfly *Drosophila melanogaster*, constitute the 'nuts-and-bolts' of developmental decision-making. The challenge to developmental biologists today is to understand the functions of these genes and to describe developmental decisions in biochemical terms. Results reported at a recent meeting* indicate that some elucidation of development at a molecular level will emerge from investigations of the nematode worm *Caenorhabditis elegans*.

Two important developments are responsible for the transition of research on *C. elegans* from traditional to molecular-genetic analysis. First, reliable methods for molecular cloning of *C. elegans* genes are now available, and second, a method of reintroducing cloned genes into the *C. elegans* germ line has now been developed. The ability to clone, manipulate and transform *C. elegans* genes offers great power and promise for unravelling genetic regulatory mechanisms.

A mechanism for determination of lineages that is often considered, but rarely explored, is chromosome imprinting (see the article by Monk on page 203). Such a mechanism predicts that parental DNA strands segregate to specific blastomeres in the embryo. The invariant cell lineage of *C. elegans* makes it a particularly attractive organism in which to investigate this mechanism. In an elegant series of experiments (K. Ito, Univ. Calgary), either oocyte or sperm DNA strands were marked in a manner that allows their segregation to be followed during embryogenesis. The cellular destination of individual chromosomes turns out to be random, a result suggesting that chromosome imprinting is not significant in early cell determination.

The cloning and molecular analysis of developmental regulatory genes was, arguably, the dominant theme of the meeting. Two general methods for cloning

genes were evident. First, techniques of 'transposon-tagging' are now well developed. Many spontaneous mutations are caused by transposition of the highly active Tc1 family of transposable elements, and such mutations allow cloning of the affected genes. In addition, two new genetically active families of transposons (J. Collins, Univ. Wisconsin; J. Yuan and M. Finney, MIT) should facilitate the cloning of additional genes. Second, an ordered set of overlapping cosmid clones, representing about 95 per cent of the *C. elegans* genome, has been assembled (A. Coulson and J. Sulston, MRC Cambridge). As the gaps in this ordered cosmid library are relatively small, chromosome walking — even over very long distances — is relatively painless. The importance of these



Photomicrograph of a *C. elegans* embryo at the four-cell stage (courtesy of J. Austin).

two techniques was apparent in many presentations. Genes representing the rich variety of *C. elegans* developmental genetics have now been isolated. Of special excitement are genes involved in the determination of cell fate (S. Kim and M. Finney, MIT; J. Way, Columbia Univ.; I. Greenwald, Princeton Univ.); sex determination (T. Rosenquist and P. Okkema, Univ. Wisconsin); programmed cell death (Yuan); developmental timing (G. Ruvkun, Harvard Univ.); and establishment of cellular asymmetry during cleavage (D. Stinchcomb, Harvard Univ.).

What is to be done with these genes when cloned? A direct and powerful approach is to manipulate them *in vitro*, to return them *in vivo* by DNA transformation and to evaluate their function in the resulting animal. Transformation and expression of five different cloned genes was described (A. Fire, Carnegie Institute, Baltimore; Way; T. Blumenthal, Indiana Univ.). Transformed genes retain their proper tissue-specific expression, but they integrate into the chromosome in

a non-homologous manner (Fire). In this respect, transformation in *C. elegans* resembles that of mammalian cells in culture.

Until recently, there was little evidence to suggest that interactions between cells are involved in early development of the nematode. The few cell interactions that were known influenced chiefly post-embryonic events. This result has been interpreted to indicate that ancestry is the primary determinant in nematode embryogenesis. At the meeting, evidence was reported for important interactions between early blastomeres. One particularly dramatic example, revealed by blastomere removal experiments, may be analogous to embryonic induction in vertebrates. An interaction between early blastomeres induces a mesodermal lineage (muscles of the anterior pharynx) from a precursor cell that primarily makes ectoderm (J. Priess, MRC Cambridge). Unexpectedly, mutations in a single gene, *glp-1*, affect both the early interaction involved in embryonic induction of the pharynx and a post-embryonic cell interaction controlling germ cell proliferation (J. Austin, Univ. Wisconsin; Priess). This finding suggests that these apparently distinct developmental events share common elements of control.

The mechanisms of gene expression in *C. elegans* hold some surprises. A compelling case was presented that trans-splicing is a regular feature of *C. elegans* messenger RNA processing (M. Krause, Fred Hutchinson Cancer Center). The 5' termini of certain mRNAs appear to be encoded in the genome as a separate leader sequence that is added post-transcriptionally during processing of the pre-mRNAs. The clearest example involves the actin gene family. Of four different nematode actin genes, three do not encode the 22 nucleotides at the 5' terminus of the mature mRNA. Rather, these nucleotides are encoded as part of an abundant, non-polyadenylated RNA that is about 100 nucleotides long. The 5'-terminal 22 nucleotides of this small RNA are attached to the actin mRNAs as an untranslated leader sequence.

The nematode phenomenon, therefore, is very similar to trans-splicing of trypanosome mRNAs. In *C. elegans*, however, not all mRNAs are trans-spliced. The fourth actin gene and the *unc-54* heavy-chain gene are not trans-spliced. The functional significance of these differences is unknown. □

Philip Anderson is in the Department of Genetics and Judith Kimble is in the Laboratory of Molecular Biology and the Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, USA.

* 1987 Meeting on *C. elegans*, Cold Spring Harbor Laboratory, New York, 6–10 May 1987.

Genomic imprinting

Memories of mother and father

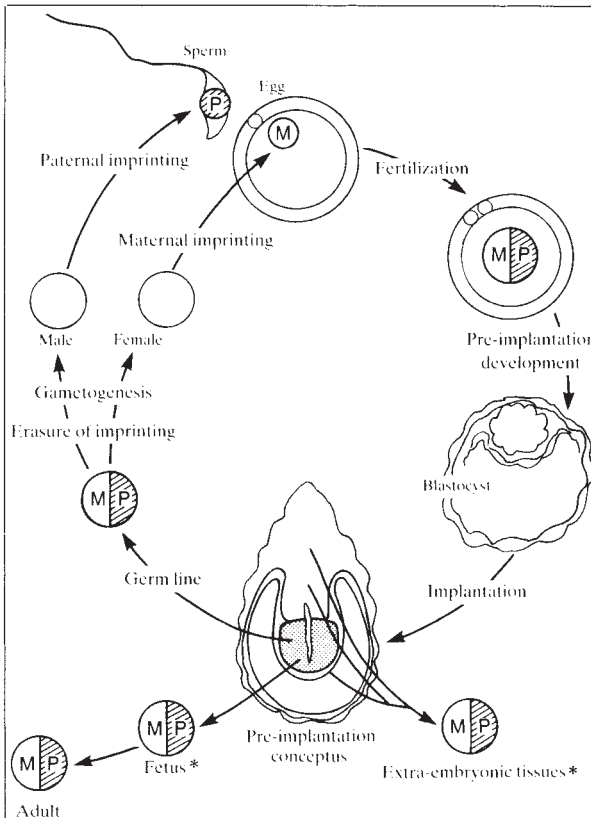
Marilyn Monk

THE maternal and paternal genomic complements inherited from egg and sperm, respectively, may remember their parental origins throughout the development and life of the individual. At fertilization, the haploid contributions from mother and father create the diploid nucleus of the individual and henceforth the maternal and paternal genes cooperate as one informational unit. Yet a memory of the gametic origin of each complement of genetic information persists, residing in some form of differential imprinting imposed on the genetic material during gametogenesis. The investigations into possible molecular mechanisms of genomic imprinting in the mouse, reported by Reik *et al.* and by Sapienza *et al.* on pages 248¹ and 251² of this issue, respectively, herald an exciting new era of discovery.

The phenomenon of differential imprinting is widespread, occurring at different levels of organization, often with significant biological consequences. For example, the inactivation and elimination of the entire set of paternal chromosomes determines maleness and male germ cells, respectively, in the mealy bug³ and in mice a genomic contribution from both the mother and the father is essential for successful development⁴⁻⁶. In addition, each genome contributes unequally in different regions of the mouse conceptus, the paternal genome being more important for development of extra-embryonic tissues, the maternal genome for embryonic development⁵. Other examples of imprinting include preferential expression of maternal or paternal genes in interspecific hybrids of sunfish⁷; preferential inactivation of the paternally inherited X chromosome in female kangaroos⁸ and in the extra-embryonic membranes of the mouse^{9,10}; and recent findings that certain segments of autosomes in the mouse must be represented by both maternal and paternal contributions for normal phenotype¹¹. No obvious connection can yet be made between these diverse observations.

Imprinting must be established during (or before) gametogenesis, persist stably throughout DNA replication and cell division in the soma, and be erased in the germ line to be differentially established once more in sperm and egg genomes (see figure). Stable, heritable, differential

modification of chromatin is required and, in mammals at least, methylation of the DNA is a possible mechanism. Differential patterns of methylation are stable and heritable and, moreover, they have been implicated in the regulation of gene expression, chromatin configuration, X-chromosome expression and imprinting



Imprinting in the mouse. M, maternal; P, paternal. Imprinting in fetus and extra-embryonic membranes (marked with asterisks) may be equivalent although 'read' differently in these different lineages. Alternatively, the original gametic imprinting may be subjected to different changes in the processes deriving the embryonic and extra-embryonic lineage during pre-implantation development leading to different modulations of the original imprinted information in these tissues.

itself¹²⁻¹⁵. Fortunately for those studying methylation, there are two isoschizomeric DNA restriction endonucleases that respond differently to methylation: *MspI* cuts the DNA at CCGG sequences whether they are methylated or not; *HpaII* cuts only the unmethylated sequence. These enzymes have been used extensively to determine the degree of DNA methylation *in toto* or in specific regions identified by hybridization to a nucleic-acid probe.

To examine the maternal and paternal DNA components of the individual separately, the authors of both reports in

this issue used the ingenious device of tagging either the maternal or paternal genome with a transgene. Transgenes integrated at different sites in the genome can show how the degree or pattern of methylation of the transgene is influenced by the genomic regions flanking the site of integration. In appropriate crosses, the transgene is inherited either from the mother or the father, and hence has been subjected to imprinting during maternal or paternal gametogenesis, respectively. Four out of five loci studied by Sapienza *et al.*², and one of seven loci studied by Reik *et al.*¹, are differentially methylated in the offspring depending on the gamete of origin. The locus studied by Reik *et al.* and one locus extensively studied by Sapienza *et al.* were both less methylated when they had been inherited from the father. (For one other locus studied by Sapienza *et al.* (data not shown) the pattern inherited from the female gamete was less methylated than that from the male gamete.) By transmitting the transgene to offspring through several generations alternatively from the mother or from the father, both groups show the imprinting methylation to be accurately and alternatively switched between maternal and paternal patterns. Whatever the mechanisms of imprinting, they are wiped out in the germ line, yet persist in the soma.

Neither of the studies reported in this issue addresses the key question of whether the methylation differences already exist in the gametes, presumably because of the difficulty in obtaining enough oocyte DNA for analysis. But, in the two cases reported here, undermethylation of the paternally inherited loci in the offspring is reflected by undermethylation of these loci in the testis. Overall, sperm DNA is more methylated than oocyte DNA (but less methylated than somatic tissue DNA) so differential methylation of the gametes themselves certainly could be the mechanism of imprinting¹⁴. This global methylation difference, however, seems to be the wrong way around for the two loci examined here. It is possible that, irrespective of global differences, imprinting resides in the differential modulation of methylation in the form of relative increases and decreases in methylation along particular domains of sperm and oocyte DNA. In somatic tissues, for example, clustered CpG sequences 5' to certain genes are more methylated on the inactive X chromosome, whereas other CpG sequences in the body of the gene are more methylated on the active chromosome (see ref. 15 for review). If imprinting is determined by

modulation of methylation patterns, then the activity of maintenance methylase could ensure that it survived through early development into the somatic tissue. Alternatively, it is possible that some other type of modification, correlated with the initial methylation differences, survives to transmit the memory.

Another question is the erasure of imprinting in the germ line. If this erasure were to be incomplete for any reason then memory would contain a grandparental component (an interesting possibility). Fetal germ cells are strikingly under-methylated¹⁴ and it is possible that the germ lineage is derived before extensive *de novo* methylation (which begins in fetal precursor cell DNA at about the time of implantation¹⁴). If there is some mechanism other than methylation levels that distinguishes maternal and paternal genomes, it must be erased in the fetal germ cells, perhaps at meiosis when possible different configurational constraints would be erased in the processes that unravel DNA before synapsis. Imprinting would be re-established during gametogenesis.

These new approaches should lead to exciting new developments in the molecular exploration of imprinting, perhaps with significant new implications for our understanding of differential gene function in development and adult life. Systematic molecular analyses will yield information on the DNA modification, chromosome configuration and patterns of transcription in maternal and paternal chromosome segments flanking the sites of integration of the transgene markers. Isolation of integration sites within segments of chromosomes known from genetic experiments to affect phenotype may elucidate the importance of the dis-

crimination between maternal or paternal origins in these regions. Segments of differential imprinting might be found which correlate with the relative importance of maternal and paternal genomes in embryonic and extra-embryonic tissues, the endogenous genes involved identified and their differential roles clarified. Finally, further studies may clarify the timing and mechanism of erasure of imprinting in the germ line. Could this event highlight an underlying process of reprogramming of the gametes to developmental totipotency?

The biological memory of mother and father built into the genetic inheritance of the new individual has profound consequences, from sex determination to survival. Maternal and paternal imprinting imposes additional information onto the actual gene sequences inherited from each parent, information that may regulate spatial and temporal gene activity as well as chromosome behaviour. How it all works remains to be seen. □

1. Brown, S.W. & Nur, V. *Science* **145**, 130-136 (1964).
2. Reik, W., Collick, A., Norris, M.L., Barton, S.C. & Surani, M.A. *Nature* **328**, 248-251 (1987).
3. Sapienza, C., Peterson, A.C., Rossant, J. & Balling, R. *Nature* **328**, 251-254 (1987).
4. Surani, M.A.H., Barton, S.C. & Norris, M.L. *Nature* **308**, 548-550 (1984).
5. Surani, M.A.H., Barton, S.C. & Norris, M.L. *Nature* **326**, 395-397 (1987).
6. McGrath, J. & Solter, D. *Cell* **37**, 179-183 (1984).
7. Whitt, G.S., Cho, P.L. & Childers, W.F.J. *exp. Zool.* **179**, 271-282 (1972).
8. Sharman, G.B. *Nature* **230**, 231-232 (1971).
9. Takagi, N. & Sasaki, M. *Nature* **256**, 640-642 (1975).
10. West, J.D., Frels, W.I., Chapman, V.M. & Papaioannou, V.E. *Cell* **12**, 873-882 (1977).
11. Cattaneach, B.M. & Kirk, M. *Nature* **315**, 496-498 (1985).
12. Holliday, R. & Pugh, J.E. *Science* **187**, 226-232 (1975).
13. Riggs, A. *Cytogenet. Cell Genet.* **14**, 9-25 (1975).
14. Monk, M., Boubelik, M. & Lehnert, S. *Development* **99**, 371-382 (1987).
15. Monk, M. *BioEssays* **4**, 204-208 (1986).

Marilyn Monk is in the MRC Mammalian Development Unit, 4 Stephenson Way, London NW1 2HE, UK.

Ionospheric physics

The sources of gravity waves

Robert D. Hunsucker

SINCE 1960, when Colin Hines wrote his seminal paper¹, *Internal Atmospheric Gravity Waves at Ionospheric Heights*, ionospheric theorists and experimentalists have published several hundred papers on this topic. Internal atmospheric gravity waves (or acoustic gravity waves, AGWs) are transverse waves that propagate in a neutral planetary atmosphere under the influence of gravity and buoyancy, and, to a lesser extent, viscosity and other forces. They should not be confused with cosmological gravitational waves. The ionospheric manifestations of AGWs are called travelling ionospheric disturbances (TIDs). Terrestrial AGWs are driven by heating by electric currents, energetic particle precipitation and 'mechanical'

forces in the auroral ionosphere, and also by volcanoes, earthquakes, tropospheric turbulence, stratospheric winds and nuclear detonations. Using optical and radio techniques, Crowley and Williams have identified a periodic source of AGW/TID phenomena, as they report on page 231 of this issue².

Gravity waves, described in several reviews³⁻⁶, are usually classified as large-, medium- or small-scale, depending on their periodicity, wavelength and horizontal velocity. Large-scale AGWs typically have periodicity, τ , of 50 minutes to 3 hours and horizontal speeds, v , of about 500 m s⁻¹ or more and seem always to originate in the auroral regions — especially during magnetic storms. These

sources are often characterized by sudden brightening of the visual aurora and intensification of the auroral electrojet (a current of about 10⁶ amperes flowing along the auroral oval). Most medium-scale AGWs ($\tau = 20-45$ min, $v = 80-450$ m s⁻¹) originate below the ionosphere, although some seem to have auroral sources. The small-scale AGWs ($\tau = 2-5$ min and $v \geq 300$ m s⁻¹) all seem to have middle- and lower-atmosphere sources associated with regions of turbulence and wind shear.

Winds and, to a lesser extent, gravity waves propagating in the neutral atmosphere, are the main processes of energy transfer from the auroral oval to the mid- and low-latitude atmosphere. Hajkowicz and I have found evidence⁷ that, at least occasionally, gravity waves and ionospheric disturbances are generated simultaneously in magnetically connected regions of the auroral oval and propagate towards the equator. Gross *et al.*⁸ suggest that the interactions of these AGWs produce stationary electron-density features in the mid-latitude and equatorial ionospheres.

Recently, experimentalists from 20 countries collaborated in a global ionospheric observation called the Worldwide Atmospheric Gravity Wave Study (WAGS) (15-18 October 1987, and followed by a workshop in January this year). The paper by Crowley and Williams in this issue² comes from WAGS, and reports a new source for AGW/TIDs. Until now, the accepted model assumed a quasi-impulse source (ionospheric joule or particle heating, Lorentz force or some other mechanism) to generate AGWs, that then produce TIDs, by collisional-coupling that propagates thousands of kilometres towards the equator. This paradigm has been verified by several investigators, most recently by Rice *et al.*⁹. Crowley and Williams report an observation of a periodic source in the auroral ionosphere that determines the periodicity of gravity waves observed an hour later propagating in the ionosphere over the United Kingdom.

More specifically, they used the European incoherent scatter radar (EISCAT) in Scandinavia to measure the joule-heating, particle-heating and Lorentz-force terms in the auroral E-region of the ionosphere and conclude that joule-heating is the dominant source. They also infer periodicities in the auroral electric field of 33 and 17 min. Because the same periodicities were observed in the waves detected in the ionosphere over the United Kingdom by high-frequency Doppler radar, Crowley and Williams conclude that the source periodicity determines the AGW/TID periodicity.

Thus, it seems that there are two distinct models of cause and effect for atmospheric and ionospheric gravity waves. In

one, a single energetic event (usually in the auroral ionosphere) generates AGWs, producing TIDs, which are then detected in the ionosphere some distance from the source. In this case, the magnitude of the source function and the properties of the terrestrial atmosphere determine the observed wave parameters. In the second case, both the magnitude and periodicity of the source function determine the observed TID parameters. It remains to be seen under what conditions the second model operates and how prevalent it is. □

1. Hines, C.O. *Can. J. Phys.* **38**, 1441–1481 (1960).
2. Crowley, G. & Williams, P.J.S. *Nature* **328**, 231–233

- (1987).
3. Yeh, K.C. & Liu, C.H. *Rev. Geophys. Space Phys.* **12**, 193–216 (1974).
4. Francis, S.H. *J. atmos. terr. Phys.* **37**, 1011–1054 (1975).
5. Hunsucker, R.D. *Rev. Geophys. Space Phys.* **20**, 293–315 (1982).
6. Fritts, D.C. *Rev. Geophys. Space Phys.* **22**, 275–308 (1984).
7. Hajkowicz, L.A. & Hunsucker, R.D. *Planet. & Space Sci.* (in the press).
8. Gross, S.H., Reber, C.A. & Huang, F. *IEEE Trans. on Geoscience Remote Sensing* **GE-22**, 340–352 (1984).
9. Rice, D.D., Hunsucker, R.D., Crowley, G. & Williams, P.J.S. *Geophys. Res. Lett.* (in the press).
10. Francis, S.H. *Bell Lab. Rep.* (1 October 1973).

Robert D. Hunsucker is at the Geophysical Institute, University of Alaska, Fairbanks, Alaska 99775, USA.

Malaria vaccines

Are they approaching reality?

Peter Perlmann

MALARIA, the most widely spread infectious disease of man, is a threat to health in areas inhabited by almost half the world's population, with an estimated number of 300 million acute cases per year and causing the death of 3–4 children every minute in Africa alone. Attempts to eradicate malaria by parasitocidal drugs and insecticides have been only partially successful and these control measures are now being seriously hampered by the rapidly growing drug-resistance of the parasite and the DDT-resistance of its mosquito vector. It is for these reasons that the development of vaccines has been considered a valid and necessary complement to control the disease. Because of the complex life cycle of the parasite and the complicated epidemiology of the disease, it has been questioned by many whether it would ever be possible to develop such vaccines. A few weeks ago, the first results were published¹ of a clinical trial of a subunit vaccine directed against the sporozoites of *Plasmodium falciparum*, the parasite species causing the most serious malaria in humans. And on page 257 of this issue², Deirdre Herrington *et al.* report the results of an independent trial run in parallel with a similar vaccine. Although it is clear from both studies that the subunit vaccines used are not the final answer, they show that protection against infection can be

achieved by vaccination.

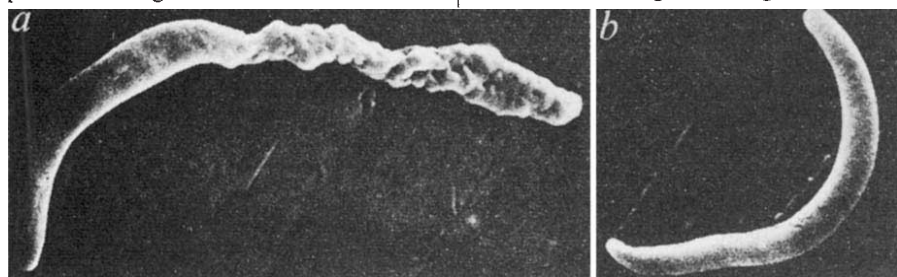
Of the three main types of malaria vaccines presently being developed³, sporozoite vaccines are directed against the life-cycle stages of the parasite that are transmitted to the vertebrate host when bitten by infected mosquitos: if efficient, these vaccines will prevent infection. Through the long-standing efforts of Ruth Nussenzweig and her colleagues at the New York University, whose latest results are reported in this issue², and work in several other laboratories, it is now well established that a significant fraction of the humoral immune response to malaria sporozoites in infected hosts is directed against a few linear epitopes of the circumsporozoite protein (CSP) that cover the surface of this stage of the parasite (see figure). Structural studies show that these epitopes are in an extensive area in the middle of this protein, an area consisting of short amino-acid sequences tandemly repeated many times. For the CSP of *P. falciparum*, the main repeat unit consists of four amino acids, asparagine, alanine, asparagine and proline (NANP in one-letter code). The vaccine used in the clinical trial of Herrington *et al.*² is a synthetic peptide, (NANP)₃, conjugated to tetanus toxoid serving as carrier. Volunteers not previously exposed to malaria were injected 3 or 4 times with different doses of the immunogen, using aluminium

hydroxide as adjuvant. The volunteers immunized with the two largest doses developed anti-NANP antibodies that also react with the parasite protein in immunofluorescence. The vaccine is safe with no adverse reactions in any of the vaccinees. When the three volunteers having the highest anti-CSP titres were subsequently challenged with sporozoites, one of them appeared to be completely protected while the appearance of blood infection in the two others was slightly but significantly delayed compared with that in four unimmunized controls.

These results are very similar to those reported recently by Ballou *et al.*¹ These authors used a vaccine prepared by recombinant DNA technology. Here the immunogen consisted of 32 four-amino-acid repeat units from the *P. falciparum* CSP and a short, 32-amino-acid-long tail encoded by the vector. In this trial, six of the immunized volunteers and two controls were challenged with sporozoites. Again, the volunteer who had the highest antibody titre was protected while the appearance of blood parasitaemia in two others, also with high titres, was delayed.

Taken together these two studies show that it is possible to induce protective antibodies in humans with subunit vaccines, using relevant antigen fragments as immunogens. The results are of importance not only for malaria, where subunit vaccines are a necessity for several reasons, but also for vaccine design in general. This does not mean that efficient sporozoite vaccines are around the corner: the problems that remain are not trivial. Here I shall discuss two of them, by no means unique to malaria vaccines.

Anti-sporozoite antibodies protect, probably by preventing entrance of the parasite into liver cells where it multiplies and develops into the stages subsequently infecting erythrocytes³. To achieve satisfactory protection, however, the antibodies raised by the vaccine must have high titres and have to be sustained sufficiently long in most of the vaccinees. With the present vaccines this is obviously not the case and there are no booster effects on repeated injections. Although some of these problems may be overcome by increasing dosage, changing the carrier system or using improved adjuvants², it is doubtful whether such changes will ever be sufficient to induce antibody-mediated protection in most vaccines, a *sine qua non* for a sporozoite vaccine of this type. The formation of anti-sporozoite antibodies requires T-cell help, now provided by the foreign carriers, tetanus toxoid in one case and, to some extent, the 32-amino acid part of the fusion protein in the other⁴. The use of foreign carriers has several disadvantages⁵ and an efficient vaccine should contain T-cell activating sites derived from the malaria sporozoite, preferentially from the CSP itself. Recent



Scanning electron micrograph of *Plasmodium* sporozoites ($\times 6,750$). *a*, Sporozoites incubated in immune serum containing anti-sporozoite antibodies develop a surface reaction termed the circumsporozoite precipitation (CSP). *b*, In contrast, sporozoites incubated in normal serum remain unaltered. (D. Herrington.)

experiments in mice indicate that this protein has only a few dominant T-helper cell activating sites which, as expected, are restricted by the major histocompatibility system (MHC class II) of the vaccine recipient^{4,6}. Moreover, although the CSP-repeat region, which in *P. falciparum* is invariant in different parasite strains, contains one such MHC restricted site, the other important T-cell sites defined thus far are in a molecular region that varies between strains⁷.

A similar genetic restriction to a few epitopes probably also holds true for the human T-cell response to the CSP (M.F. Good *et al.*, manuscript in preparation). The findings imply that it may be necessary to use a vaccine formulation incorporating T-cell activating structures derived from other sporozoite molecules than the CSP to achieve the goals indicated above. This is by no means impossible⁸ but remains to be investigated.

The second major problem is a related one and also involves T-cell activation. Recent animal experiments suggest that an important part of the protective effect of a sporozoite vaccine is antibody independent⁹, probably reflecting a direct T-cell-mediated inhibition of parasite development in the liver¹⁰. If true for human malaria as well — and there is no

reason to believe that it is not — this also means that one will probably have to search for additional T-cell activating sporozoite molecules to design an efficient subunit vaccine.

Despite the problems discussed above (and others that I have not mentioned here), the reports by Ballou *et al.*¹ and Herrington *et al.*² are important steps on the way to the design of vaccines preventing malaria infection. Although the remaining problems are serious, they are not insurmountable. Considering the fact that it took only 3 years from the structural definition and cloning of *P. falciparum* CSP until the clinical trial of the first vaccines, it may well be hoped that the second or third generation of improved sporozoite vaccines will become available in the foreseeable future. □

1. Ballou, W.R. *et al. Lancet* **i**, 1277-1281 (1987).
2. Herrington, D.A. *et al. Nature* **328**, 257-259 (1987).
3. Miller, L.H. *et al. Science* **234**, 1349-1356 (1986).
4. Good, M.F. *et al. J. exp. Med.* **164**, 655-660 (1986).
5. Herzenberg, L.A., Tokuhio, T. & Herzenberg, L. *Nature* **285**, 665-667 (1980).
6. Del Giudice, G. *et al. J. Immunol.* **137**, 2952-2955 (1986).
7. Good, M.F. *et al. Science* **235**, 1059-1062 (1987).
8. Scherle, P.A. & Gerhard, W. *J. exp. Med.* **164**, 1114-1128 (1987).
9. Egan, J.E. *et al. Science* **236**, 453-456 (1986).
10. Ferreira, A. *et al. Science* **232**, 881-884 (1986).

Peter Perlmann is in the Department of Immunology, University of Stockholm, S106 91 Stockholm, Sweden.

Polymer science

Interdiffusion at interfaces

Barton A. Smith

Two pieces of plastic can be welded together by bringing them into contact at an elevated temperature. This process has many practical applications, including the joining of plastic pipes, the moulding of plastic parts from powdered resins, and adhesion in multilayered materials. The strength of the weld results from the interdiffusion of the polymer molecules across the initial interface. The interdiffusion process for polymers differs from that in small-molecule and atomic solid systems, because of the entangled state of the long polymer chains in the melt. There has recently been considerable controversy over the type of theory which correctly describes polymer-polymer interdiffusion¹. Proposed theories have differed markedly in their predictions of the way in which the rate of interdiffusion depends upon parameters such as the length of the polymer molecules. Composto *et al.* describe on page 234 of this issue² some experiments carefully designed to provide a direct test of these theories.

In the simplest case, the welding of identical materials, the interdiffusion proceeds at the same rate as diffusion of those molecules in the bulk sample, and is characterized by the self-diffusion

coefficient. Self diffusion has been studied extensively over the past decade and is described adequately by the 'reptation' model of polymer motion. The more interesting cases involve the interdiffusion of different polymers, or chemically identical polymers of very different molecular weights. These cases are more complex because the mixing is driven by both entropy and enthalpy, and because both the mixing energy and the mobility of each molecule depend upon the composition of the material, which varies continuously across the interface.

Attempts have been made to predict the mutual diffusion coefficient, D_M , which characterizes the interdiffusion rate, from the tracer or self diffusion coefficients of the individual components. (The tracer diffusion coefficient describes the random brownian motion of individual molecules, in the absence of any composition gradients.) As the general case is too complex to solve analytically, theories have been based on certain simplifying assumptions and boundary conditions. The recent controversy has centred on two assumptions that lead to different forms for the equation predicting D_M .

One theory³ allows for a flux of

vacancies across the interface and assumes that the osmotic pressure is constant, resulting in D_M being a linear combination of the mobilities of the two components. Thus, D_M is controlled primarily by the mobility of the more rapidly diffusing polymer, in agreement with experimental results. The vacancy flux model has been criticized because it allows a significant density gradient across the interface.

The alternative assumption, that the density is constant, requires the fluxes of the two components to be equal and opposite. The result is that the mobilities of the two components are strongly coupled. A 'slow' theory⁴ resulted from the idea that the tracer diffusion coefficient of the slow component would control D_M . Recent experiments would seem to confirm the vacancy flux theory. For example, Composto *et al.*² describe in this issue a series of measurements of both the tracer and mutual diffusion coefficients in the miscible blend of polystyrene and poly(xylenyl ether). They find that D_M scales inversely with the molecular weight of the polystyrene, the faster moving species, and report good quantitative agreement between their data and the predictions of the vacancy flux theory. The 'slow' theory is clearly in error, predicting the interdiffusion to be more than ten times slower than the measured rate in some cases.

Recently it has been realized, however, that the constant-density assumption does not necessarily lead to a 'slow' theory. The resulting mutual diffusion (for each composition) is analogous to that obtained for a dilute or semi-dilute polymer solution: D_M depends strongly upon the mobility of the solvent, which in this case is the more rapidly diffusing polymer. This has also been described as the convective flow of the slow molecules caused by the influx of the rapidly diffusing ones⁵. Jordan *et al.*⁶ have recently analysed the concentration profiles of the interdiffusion of different molecular-weight polyethylenes. They confirm that the faster component controls D_M , and suggest that the current thinking along the lines of the constant density model may lead to the best description of polymer-polymer interdiffusion. We can look forward to further experiments of this type to guide us to a more detailed theoretical description of the process. □

1. Brochard-Wyart, F. in *Fundamentals of Adhesion* (ed. Lee, X.) CRC, Cleveland, in the press.
2. Composto, R.J., Kramer, E.J. & White, D.M. *Nature* **328**, 234-236 (1987).
3. Kramer, E.J., Green, P. & Palmström, C.J. *Polymer* **25**, 473-480 (1984).
4. Brochard, F., Jouffroy, J. & Levinson, P. *Macromolecules* **16**, 1638-1643 (1983).
5. Sillescu, H. *Makromol. Chem. Rapid Comm.* **5**, 599-603 (1984).
6. Jordan, E.A. *et al. Macromolecules* (in the press).

Barton A. Smith is at the IBM Almaden Research Center, San Jose, California 95120, USA.

A relativistic jet from SN1987A?

SIR—Speckle interferometry of supernova SN1987A 5–7 weeks after the explosion has revealed a second object only three magnitudes fainter than the supernova itself, with an angular separation of 0.06 arc s (refs 1,2). The corresponding projected distance is around 5×10^{16} cm, implying a velocity of at least $0.3c$ if it is triggered by the supernova. I interpret this observation as evidence for a relativistic jet, and discuss here how a central pulsar might give rise to such a jet.

Newly formed pulsars may spin much faster than the Crab pulsar and release far more power than the 10^{38} erg s^{-1} now being injected into the Crab nebula. During the first few months after its birth, a pulsar would be shrouded by the expanding supernova envelope; but its power output may modify (or even dominate) the light curve^{3–7}.

The Crab pulsar generates more power in the form of a relativistic wind^{4,8} than as actual pulsed radiation, and this may also be the case for newly born pulsars. This plasma, being coupled to the magnetic field spun off the pulsar, has the attributes of a fluid. Unlike the radiation, it cannot escape by diffusing through the remnant. Were spherical symmetry strictly maintained, the relativistic plasma would gradually inflate a bubble within the remnant, which would itself be gradually accelerated by the build-up of pressure inside it. But if the remnant were 'punctured', relativistic plasma from the over-pressured bubble would squirt into the surrounding interstellar medium. Once the plasma had an escape route, essentially all of the pulsar's power output would be channelled into a jet.

How might the remnant have developed a leak? Perhaps simply via a Rayleigh–Taylor instability. But a more efficient mechanism suggests itself if the star that exploded had a close binary companion: some ejecta from the explosion would collide with this companion; the expanding remnant would consequently lose spherical symmetry, being underdense in the directions shadowed by the companion; when the companion had moved away, relativistic plasma could flow out from within this cone.

After 'breaking out' from the remnant (which itself expands at $\sim c/30$), the situation resembles the jets in extragalactic radio sources⁹: relativistic fluid carrying an energy flux L_{rel} would advance at a speed V controlled by the balance between momentum density and the ram pressure $\rho_{\text{ext}} V^2$ of the external medium (with density ρ_{ext}). This speed would be $\sim c$ until the momentum density dropped to $\rho_{\text{ext}} c^2$ this happens when the jet penetrates to a radius r_0 such that $L_{\text{rel}}/r_0^2 \Omega c^3 =$

ρ_{ext} , where Ω is the solid angle. In the present context it is convenient to write $r_0 = 5 \times 10^{16} (L_{\text{rel}}/10^{41} \text{ erg } s^{-1})^{1/2} (\rho_{\text{ext}}/10^{-24} \text{ g } \text{cm}^{-3})^{-1/2} \Omega^{-1/2} \text{ cm}$. For $r > r_0$ a bow shock would advance at a speed $V \propto L_{\text{rel}}^{1/2} \Omega^{-1/2} r^{-1}$. If L and Ω were constant, the velocity would be proportional to r^{-1} , and the jet would lengthen as $t^{1/2}$. The realistic dynamics would be more complicated (neither L_{rel} nor Ω would be constant) but these considerations suggest that, for plausible parameters, the shock might form at $r_0 \approx 10^{16}$ – 10^{17} cm, a few weeks after the plasma 'breaks out', and thereafter advance at a near-relativistic (but decelerating) speed.

Where the jet runs into the external medium, its bulk kinetic energy converts into relativistic particles at a shock or 'working surface'. The equipartition magnetic field strength is $\sim (L_{\text{rel}}/10^{41} \text{ erg } s^{-1})^{1/2} (r/10^{16} \text{ cm})^{-1} \Omega^{-1/2}$ Gauss; the electrons radiating in the optical band (and at all higher frequencies) would, for any fields ≥ 0.1 Gauss, decay in less than a week. The synchrotron luminosity from the jet therefore depends on the efficiency with which the jet's kinetic energy gets converted into relativistic electrons.

Provided that an 'escape cone' exists, most of the pulsar's power output would be channelled along it, and much of this could be radiated by a jet-like feature. The jet luminosity could well be comparable with that of the supernova envelope itself. Direct synchrotron radiation offers the most obvious mechanism, in which case the jet may be strongly polarized and detectable in wavebands other than the optical. An additional possibility is that cool material from the supernova envelope, entrained by the jet, is compressed by the high pressure to densities 10^{10} cm^{-3} . At this density, only 10^{23} g of gas, excited by synchrotron radiation or by direct dissipation of kinetic energy, could emit as much as 10^{40} erg s^{-1} in (perhaps Doppler-shifted) lines. Non-thermal radiation emitted at infrared and longer wavelengths could be absorbed and reprocessed by this gas. (Scattering or reprocessing by a cloud at ambient interstellar densities cannot achieve high enough volume emissivity; this is the problem with a straightforward 'light echo' interpretation of the observed feature in SN1987A.)

The scale and luminosity of the speckle feature are entirely consistent with it being a jet. SN1987A must then harbour a fast pulsar powerful enough to affect the light curve; much of this power could now be escaping as relativistic plasma. The optical feature revealed by speckle analysis corresponds to the place where the jet is running into the interstellar medium. In wavebands other than the optical, this feature could conceivably outshine the supernova itself. Future studies of its motion, polarization and spectrum should

help delineate the dynamics and evolution of the putative jet.

I thank Andy Fabian, George Field, Bob Kirschner and Peter Meikle for helpful discussions.

MARTIN J. REES

*Institute of Astronomy,
Madingley Road,
Cambridge, CB3 0HA, UK
and
Center for Astrophysics
60, Garden Street,
Cambridge,
Massachusetts 02138, USA*

1. Karovska, M., Nisenson, P., Noyes, R. & Papaliolos, C. *IAU Circ.* No. 4382 (1987).
2. Marcher, S.J., Meikle, W.P.S. & Morgan, B.L. *IAU Circ.* No. 4391 (1987).
3. Rees, M.J. *Astrophys. Lett.* **6**, 55 (1970).
4. Ostriker, J.P. & Gunn, J.E. *Astrophys. J.* **184**, L95 (1971).
5. Bisnovatyi-Kogan, G.S. *IAU Symp.* No. 66, 152 (1974).
6. McCray, R.A., Shull, M. & Sutherland, P. *Astrophys. J.* (in the press).
7. Ostriker, J.P. *Nature* **327**, 287 (1987).
8. Rees, M.J. & Gunn, J.E. *Mon. Not. R. astr. Soc.* **167**, 1 (1974).
9. Begelman, M.C., Blandford, R.D. & Rees, M.J. *Rev. mod. Phys.* **56**, 255 (1984).

Why did Nature select green plants?

SIR—Terrestrial plants look green because chlorophyll absorbs light at the red and blue ends of the spectrum, with the central region (the green) being reflected or transmitted. As an ideal photosynthetic pigment should be black — absorbing at all visible wavelengths — why did nature choose a green one?

A possible answer lies in the photosynthetic bacterium *Halobacterium halobium*. This organism lives in salt lakes and bears all the hallmarks of being a living fossil, surviving only because it can tolerate extremely high salt concentrations where other organisms cannot compete. It has a membrane structure not found in other present-day organisms with terpenoids in place of the fatty acids¹ (similar terpenoids are found in ancient rocks² and may pre-date fatty acids as membrane components). Instead of chlorophyll, it has the purple pigment bacteriorhodopsin, the chromophore of which is the same as that of rhodopsin, the light-sensing pigment in the eyes of animals¹. This type of organism probably evolved before animals and plants: from such organisms, plants inherited their visual pigment and plants their carotene.

Photosynthesis with bacteriorhodopsin is very primitive. The light-excited pigment transports protons across the photosynthetic membrane: the resulting gradient is used for the active uptake of nutrients, locomotion and the chemiosmotic production of ATP. There is no evolution of oxygen or fixation of carbon dioxide. Such bacteria, which still need a supply of organic materials to grow, would probably have been well-adapted to the

illuminated 'primaevial broth' at the dawn of life. Since they were possibly the first photosynthetic organisms, the whole of the spectrum not absorbed by water (from the edge of the ultraviolet to the near infrared) was available to them, which is probably why natural selection gave them the photo-synthetic pigment bacteriorhodopsin, with its broad absorption peak in the middle of this region.

Chlorophyll-based photosynthesis is much more complex. Bacteriorhodopsin uses light to transport protons across a membrane. Chlorophyll transports electrons instead. It too generates a proton gradient for ATP production, but does so indirectly as a result of relatively complex electron-processing reactions on either side of the photosynthetic membrane. To avoid the need to replicate these systems for each chlorophyll molecule, most of the pigment molecules are arranged in organized arrays around specialized molecules in reaction centres to which they pass their light energy by resonance transfer¹. Chlorophyll-based photosynthesis almost certainly evolved after bacteriorhodopsin-containing organisms were well established, when new photosynthetic organisms stood the best chance of survival if their photo-synthetic pigments did not absorb in the same region as bacteriorhodopsin. It is probably for this reason that nature selected the green pigment chlorophyll *a*, with its absorption peaks on either side of the rhodopsin peak.

The main advantage of chlorophyll is that its light-driven electron transport can drive redox systems, in particular the reduction of CO₂. This ability almost certainly enabled the early chlorophyll-containing organisms to take over from those with bacteriorhodopsin as the primaevial broth became depleted of organic nutrients. With the near-extinction of the rhodopsin-containing bacteria, more light would have become available in the centre of the spectrum where chlorophyll *a* does not absorb, allowing the evolution of a whole series of accessory pigments that absorb in this region and which pass on their energy to chlorophyll by resonance transfer.

Algae living in poor light in deep water have very effective accessory pigments. The combination of chlorophyll with phycoerythrin in the red algae covers a large part of the spectrum and the thallus of these organisms tends to be very dark in colour, approaching black, the 'ideal'. Similarly, the fucoxanthol of brown algae and diatoms complements chlorophyll very well. These pigments very efficiently pass their energy to the reaction centre³, but bear very little resemblance to chlorophyll, suggesting that they may have arrived by a long evolutionary pathway.

Where selection pressure was less intense, less drastic solutions are found.

Land plants, and the green algae living near the surface, are not short of light and a minor modification to some of their chlorophyll molecules to give chlorophyll *b* was sufficient. This brought the twin absorption peaks of the pigment closer towards the centre of the spectrum to absorb light missed by chlorophyll *a*, but there is still a gap in the middle where light can escape unabsorbed, giving these plants their characteristic green colour.

ANDREW GOLDSWORTHY

Department of Pure and Applied Biology,
Imperial College,
London SW7 2BB, UK

1. Schreckenbach, T.H. in *Phyosynthesis in Relation to Model Systems* (ed. Barber, J.) 189–209 (Elsevier, Amsterdam).
2. Calvin, M. *Chemical Evolution* (Clarendon, Oxford, 1969).
3. Hall, D.O. & Rao, K.K. *Photosynthesis* (Edward Arnold, London, 1987).

Prospects of infrared prospecting for planets

SIR—Interest in the search for planets around nearby stars has recently been fuelled by apparent evidence from the Infrared Astronomical Satellite (IRAS) of proto-planetary material orbiting Vega¹ and possibly many nearby main-sequence stars². The discovery of what appears to be a disk viewed edge-on in β Pictoris³ has given "Prospecting for planets in circumstellar dust"⁴ a further boost. While the presence of even a little dust around a solar-type star may be suggestive, unfortunately it makes any search for embedded planets a formidable task, at least in the infrared.

This can be illustrated by calculating what our Sun and its surrounding zodiacal dust look like in the infrared when viewed from 10 parsec (pc) away in comparison to some of the Vega-like stars found by IRAS and to Earth- and Jupiter-size planets. Ten pc is taken as a representative distance known to contain about 50 solar-type stars. Good *et al.*⁵ have modelled the zodiacal cloud using IRAS data with a density

$$r(R, T) = er_0 (R_0/R)^{1.1} \exp[-4.2(Z/R)^{1.2}],$$

where $R_0 = 1$ AU, $er_0 = 1.1 \times 10^{-20}$ cm⁻¹ between $R=0.1$ and 5 AU, zero elsewhere and the temperature distribution $T(R) = 282(R_0/R)^{0.5}$. The table gives the calculated infrared flux from the Sun, the zodiacal dust, Jupiter and Earth, as seen from a distance of 10 pc. The Sun, Jupiter and Earth are taken to be 5,800 K, 135 K

and 280 K black bodies, respectively. Results are given for 1 and 0.05 arc s beam diameters. A one arc s diameter beam (equal to the diffraction limit for a 2-m telescope at 12 μ m) is just large enough to enclose the entire Solar System out to the orbit of Jupiter. With a 0.05 arc s beam, equivalent to 0.5 AU at 10 pc and requiring a 40-m telescope (array) in space, the telescope can be pointed as close as 1 AU from the star without the warmest zodiacal dust or direct starlight entering the beam.

The table shows that the zodiacal flux in the one arc s diameter beam exceeds the flux from Jupiter by more than a factor of ten, but is only 0.002 of the photospheric flux from the Sun at 100 μ m and even less at 60, 25 and 12 μ m. Given that the uncertainty of photospheric flux models of solar-type stars is in the several per cent class, it is clear that evidence of zodiacal dust, not to mention any planets, would not be detectable in the IRAS data or with a 2-m class telescope. For nearby stars identified as Vega-like from IRAS data, the flux from the circumstellar disk exceeds the flux from the central star by factors 2.5–146 in the extreme case of β Pic². This is greater than three orders of magnitude more than the radiation due to zodiacal dust in our Solar System.

The flux from the zodiacal light with a 0.5-AU diameter beam pointed a distance of 1 AU from the Sun beam would be about 60 per cent larger than the signal from an Earth-size planet. This ratio is encouraging. But the search for any embedded planets will be frustrated if the density distribution of the dust is not smooth and the diffracted light from the star is not almost totally suppressed.

I conclude that prospects for remote prospecting in the infrared for Earth and Jupiter-size planets in Vega-like circumstellar disks with any size telescope are virtually nil. Prospects are more encouraging for stars with dustclouds no larger than our zodiacal cloud. But that would require a telescope an order-of-magnitude larger than will be built in the foreseeable future.

HARTMUT H. AUMANN

Jet Propulsion Laboratory,
California Institute of Technology,
Pasadena, California 91109, USA

1. Aumann, H.H. *et al. Astrophys. J.* **279**, L23 (1984).
2. Aumann, H.H. *Publs. astr. Soc. Pacif.* **97**, 885 (1985).
3. Smith, B.A. & Terrell, R.J. *Science* **226**, 1421 (1984).
4. Diner, J.D. & Appleby, J.F. *Nature* **322**, 436 (1986).
5. Good, J.C., Hauser, M.G. & Gautier, T.N. *Advances in Space Research* 1987, (in the press).

Flux in the beam [Jy]

micrometre	10-AU diameter beam Sun centred			0.5-AU diameter beam at 1AU from Sun		
	Sun	Zodiacal	Jupiter	Earth	Zodiacal	
12	1.73	1.3×10^{-4}	5.8×10^{-7}	4.6×10^{-7}	7.3×10^{-7}	
25	0.42	1.8×10^{-4}	6.4×10^{-6}	5.2×10^{-7}	8.5×10^{-7}	
60	0.07	1.0×10^{-4}	6.7×10^{-6}	1.9×10^{-7}	3.1×10^{-7}	
100	0.027	5.3×10^{-5}	3.7×10^{-6}	8.3×10^{-8}	1.4×10^{-7}	

Breaking the chain

Ian Smart

A Changing Climate: Environmentalism and its Impact on the European Energy Industries. By Andrew Holmes. *Financial Times Business Information**:1987. Pp.176. Spiral bound £157, \$251.

Acid Deposition and Vehicle Emissions: European Environmental Pressures on Britain. By Peter Brackley. *Gower*†: 1987. Pp.136. £19.50, \$36.50.

BOTH OF these essays start ostensibly from a common premise: that organized concern for the natural environment has come to be a major constraint on the energy systems of Britain and its European neighbours. The point is explicit in the title and content of Holmes's work, but equally clear, if only implicit, in Brackley's (which in any case comes from the Joint Energy Programme of the Policy Studies Institute and the Royal Institute of International Affairs). Each describes how arguments about acid rain bear on the design and management of electricity supply. Each recognizes that worries about car exhausts are reflected in the oil companies' plans and prices. So presumably is the cost of oil pollution at sea, with which Holmes also deals. And there can surely be no clearer case of environmental pressure on energy development than that of nuclear power, for which he reserves his longest chapter.

The original proposition seems all the more firmly founded because the link between energy and the environment is two-way. Holmes argues persuasively that "energy has become central to environmental activity" (p. 2). In his foreword to Brackley's study, Jonathan Stern refers to "the growing importance of environmental issues in ... energy decision-making" (p. vii). The two studies appear moreover to reinforce each other by adopting complementary emphases: Brackley is principally interested in the activities of governmental and industrial bodies, while Holmes concentrates more on environmental pressure groups. Taken together, the two should thus provide powerful support for the initial premise. Why, then, is the effect of reading them a deepening sense of unease about the subject as a whole?

Lesser reasons for disquiet lie in the uneven qualities of the particular investigations. Brackley offers a fund of carefully organized and referenced information, including an exemplary technical appendix for the layman. But the fact that almost all of his material is arranged under national sub-headings makes it unnecessarily difficult to trace thematic threads through the undergrowth. More importantly, he has

surprisingly little to say about the larger political, economic or even energy significance of the two environmental issues he has examined. It is almost as though they matter because they are on the international agenda, rather than vice versa.

In contrast, Holmes, after a selective introduction to the environmental movement, spends much of his time on a some-

"The road hedged by such 'political realities' may conceivably begin with burning cleaner coal, but it ends with burning books and witches."

what rambling scrapbook of notes on the histories of four topics: marine pollution by oil tankers, vehicle emissions and lead in motor fuel, acid rain, and the opposition to nuclear power. One might look for as much in a series of feature articles. These rather undigested case studies are sandwiched, however, between two short chapters, "Making an Issue of It" and "Into the 1990s", in which a number of the real questions about environmental pressure on energy policy are at last identified. The few pages concerned are well worth thinking about, even if they cannot entirely make up for slim pickings elsewhere.

The sources of more serious unease relate to just those fundamental questions, and exist on two deeper levels. Holmes takes us to the first level by recalling Ashby's classic description, in *Reconciling Man with the Environment* (1978), of the three-stage chain reaction which can convert recognition of an environmental hazard into political action to control it: a reaction leading from mobilizing and igniting public opinion, through objective evaluation of the risk, to a policy founded on both information and judgement. On that level, the apparent complementarity of these two volumes turns out to mark a considerable disagreement between their authors, especially about the 'ignition' stage.

Holmes is in no doubt that the spark to set public opinion alight comes from voluntary groups of environmental activists. Symptomatically, his appended directory of environmental organizations lists numerous European Green parties

and branches of Friends of the Earth (although for some reason not of Greenpeace), but makes no mention of governmental and intergovernmental agencies or mere traditional conservationists. Brackley, however, has less time for what he refers to in the British case as "strident and obvious environmental lobbies" (p. 74). Even in West Germany, he concludes that "TV programmes and weekly magazines ... have been much more influential than any environmental movement" (p. 57). Instead, he evidently believes that what make things happen are conventional politics, operating nationally through parliamentary assemblies and internationally through such bodies as the EEC, the OECD and the UN. And his own appendix on environmental institutions exactly reflects that opposite conviction.

This might seem no more than a difference of judgement about an inevitably obscure and changeable process. To suppose that, however, would be to miss the deepest questions of all. For what ultimately divides the two books and bedevils the contemporary debate with which they are variously concerned is a profound uncertainty about the true nature of issues and choices labelled as environmental, the true aim of policies advocated to resolve them, and the true bearing of those issues and policies on the energy business.

Taking the last uncertain element first, it is notable that Brackley, despite his study's genesis, hardly explores the question of how fuel suppliers and energy converters will be affected by measures to clean up emissions from power stations and vehicles. It is even more surprising that Holmes, despite his subtitle, does little to repair the deficiency. He asserts the importance of marine pollution to governments and shipowners, but offers no guidance to its effect, if any, on oil companies or other groups in the energy sector. He says something about energy industry responses to moves towards lead-free petrol, but makes only a perfunctory attempt to evaluate the costs and benefits involved, for example to refiners. He describes the development of standards for SO₂ and NO_x emissions, but fails to assess their prospective effect on fuel choices and other planning decisions by the electricity supply industry.

Only when he comes to nuclear power, where his own feelings are evidently most engaged, does Holmes commit himself with any certainty to a conclusion about the impact of environmentalism on the relevant European energy industry. Although it represents only one of the obstacles to nuclear electricity, the environmental movement, he says, has catalysed a public and political reaction from which it is now unlikely that the nuclear power industry can recover in the present

*102 - 108 Clerkenwell Road, London EC1M 5SA, UK. †Gower House, Croft Road, Aldershot GU11 3HR, UK/Old Post Road, Brookfield, Vermont 05036, USA.



Pressure groups — Brokdorf, West Germany, June 1987. Disguised demonstrators throw stones in front of a nuclear power station during an anti-nuclear demonstration. Does the future of environmental change lie with voluntary activists or the conventional politics of parliamentary assemblies and international bodies?

century. Indeed, it may not survive.

The irony is that it is this last case of nuclear power which also arouses the greatest uncertainty about the very meaning of 'environmental'. The problem has always been there in the energy context. During the 1970s, what appeared to be fears about the environmental effect of burning fossil fuels sometimes turned out to be worries about resource depletion rather than ecological damage. More often still, there is confusion between the influence of energy activity on the natural environment and its impact on the health and safety of human individuals.

But the nuclear debate, as Holmes and many others interpret it, has extended the meaning of environmentalism in a more dramatic and questionable manner. To use his own words, it has become not an issue of energy and the natural world but "a debate about the future of industrial society" (p. 107). It is thus the extreme case of a tendency which he identifies as distinguishing the environmentalist movement born in the 1970s from the middle-aged, middle-class nature societies of a previous vintage. The new environmentalism, with its characteristically "low-growth, decentralised approach" (p. 144), embodies "opposition to the whole of industrial society" (p. 5), and its attitude to energy policy is a function of that ideological 'world view' rather than the expression of some specific ecological concern for its own sake.

There can be no possible objection to pressure groups openly opposing the

existing social and economic order, provided only that their tactics are democratically acceptable. But the significance of their advance under an allegedly environmental banner is not to be minimized. Holmes has described a process involving not only the increasing political influence of environmentalist groups but also their own progressive politicization. (His subsequent assertion that environmentalist groups other than the Green parties have maintained a non-political approach is directly contradicted by his own evidence.) And the paradoxical effect of politicization has been to drive a wedge between many avowed environmentalists and the environment itself, in that the former come to see protection or improvement of the latter as incidental to pursuing a larger political goal.

One result is that it becomes possible to resolve the apparent disagreement between Holmes and Brackley over which organizations play the main role in promoting environmental policies to do with energy. Each is right, because each has his separate idea of policy's means and ends. The stately development of international and national regulation which Brackley describes genuinely has environmental quality as an end, whereas the groups which preoccupy Holmes really regard both energy and ecology as second-order issues. Those responsible for energy decisions in government or industry must nevertheless recognize the influence of Holmes's new-style environmentalism, as well as its often passionate sincerity. But their responses have to take account of the fact that environmentalists in this category are commonly looking beyond the ostensibly ecological agenda.

There remains a final, more sinister implication, generated by the way politicians have sometimes been inclined to react to the new-style environmentalism. Environmental policies, including those with repercussions on energy, must after all be determined by political methods. But that does not permit those who formulate them to be ruled by emotion or expediency, divorced from factual evidence. Brackley sees the point when he writes: "Decisions will always be taken politically, but the quality and the effects of decisions would be enhanced if all were better informed (p. 75).

Both of these authors, however, identify all too clearly instances where the politicization of the environmental movement and its success in raising the political stakes and political temperature have induced responsible decision-makers to succumb to pressure in default or in spite of the facts. Thus Holmes correctly remarks that the British government's 1983 decision to eliminate lead from petrol, despite countervailing scientific doubt, showed that "political pressure counts for more than evidence" (p. 26) — as it did

again in 1987, when the government precipitately changed its policy on low-level nuclear waste. And Brackley recalls that a European Community commissioner in 1985 defended a decision to ban hormones in meat production on the grounds that political realities deserved more attention than scientific facts.

The importance of that shift from conviction to coercion cannot be over-emphasized. It is as much as to eliminate the crucial middle stage from Ashby's celebrated chain reaction, leaving no intervention of objective inquiry between an ignition of popular sentiment and an environmental policy decision. As such, it represents a far more serious threat to both industry and society than any ban on sulphur emission or nuclear fission. The road hedged by such 'political realities' may conceivably begin with burning cleaner coal, but it ends with burning books and witches. □

Ian Smart, 3 Grosvenor Avenue, Richmond, Surrey TW10 6PD, UK, is an independent adviser on energy affairs.

Universal physics

Robert M. Wald

Gravity, Gauge Theories and Quantum Cosmology. By Jayant V. Narlikar and T. Padmanabhan. Reidel: 1986. Pp.468. \$97, Dfl.220.

IN THE past 15 years there have been many developments in particle physics, cosmology, quantum gravity and related topics. In this book, the authors have the ambitious goal of summarizing these developments as well as presenting their own approach to quantum cosmology.

The volume is divided into four parts. The first introduces quantum field theory and gauge theories, while the second gives an introductory treatment of classical general relativity. Part III discusses quantum effects occurring in a classical space-time geometry. Finally, Part IV contains a brief account of quantum gravity followed by several chapters on the authors' own work in quantum cosmology. The book is intended to be introductory, and I would judge it to be pitched at the level of an advanced graduate student in physics. Very few references to the original literature are given.

Any book of this scope is certain to have drawbacks. In this case, it is hard to imagine a reader with little or no background in quantum field theory and general relativity who could learn enough in reading Parts I and II to be able to follow Parts III and IV. On the other hand, a researcher with an adequate background in those subjects probably would not find much of use in the introductory parts (which, together with the appendices, take up

three-quarters of the book). However, some readers could find inspiration in the wide range of exciting topics covered, and one has the impression that they are the main intended audience.

A problem with the balance of material occurs in Part IV. The summary of all approaches to quantum gravity given in Chapter 11 (of only 22 pages) is extremely sketchy. Yet two chapters plus an epilogue (86 pages) are devoted to the authors' own work in quantum cosmology. Their approach consists of treating the conformal geometry of spacetime classically and viewing the conformal factor as a quantum field. Most researchers would not accept this rather *ad hoc* hypothesis and would raise a number of technical objections. Of course, this same criticism probably applies to all other current approaches to quantum cosmology, and there is some merit in the authors' comment in the preface that they should be entitled to their share of prejudice. Nevertheless, in a book apparently aimed primarily at students it seems to me inappropriate to elaborate in detail on speculations resulting from one approach while barely mentioning results obtained from others.

The weakest sections of the book are those in which the authors deal with abstract mathematical material. I found these discussions difficult to follow and encountered several glaring errors. (For example, on p. 114 the definitions of "simple" and "semi-simple" Lie algebras are not correct, and it then seems to be argued that all finite dimensional representations of an arbitrary Lie group can be assumed to be unitary.) The dismissive attitude towards mathematical material — expressed in statements such as "Those with a healthy suspicion for abstract topics like 'homotopy classes' . . ." — does not increase one's confidence that the authors consider these issues sufficiently important to take the care needed to get all the details correct.

The quality of the coverage of the less abstract and formal topics is much higher. Although sketchy in places, the derivations and discussions appear generally to be sound. However, even here there are occasional misstatements and errors.

The main strength of the book, in my opinion, lies in the numerous examples (138 in total) given in "fine print" throughout the text. These consist of model problems, intermediate steps of derivations and other points which would normally be omitted in textbooks. In general, enough detail is given to enable the reader to work through the examples, and I believe that they are potentially a quite useful learning aid. □

Robert M. Wald is a Professor in the Physics Department and Enrico Fermi Institute, University of Chicago, 5630 Ellis Avenue, Chicago, Illinois 60637, USA.

An arresting phenomenon

John Guckenheimer

When Time Breaks Down: The Three-Dimensional Dynamics of Electrochemical Waves and Cardiac Arrhythmias. By Arthur T. Winfree. Princeton University Press: 1987. Pp. 339. Hbk \$60, £37.70; pbk \$19.95, £12.50.

PHASE is incompatible with time. This statement is an interpretation of the basic topological fact that there is no way of defining a continuous, increasing function on a circle. Arthur Winfree has devoted his career to exploring the biological consequences of this theorem and analogous topological results in two and three dimensions. *When Time Breaks Down* is the third book he has written summarizing his work. More than the earlier two (*The Geometry of Biological Time* and *The Timing of Biological Clocks**), the emphasis here is upon two- and three-dimensional phenomena. The main topics under discussion are spiral and scroll-shaped waves as they occur in chemical systems and cardiac arrhythmias. Whether or not we find here the key to understanding sudden cardiac death and fibrillation, the book is fascinating reading — it helps us to stretch our imagination in thinking about biological problems. But although I am generally enthusiastic about the book, I do have some reservations to which I shall come later.

Winfree writes in a very personal style, and casts himself in the role of a detective whose mentor is Sherlock Holmes. He is working on the case of sudden cardiac arrest and trying to nab the infamous scroll ring. The accused has not come forward with an admission of guilt, so the mystery does not end with the definitive answer that one might expect. In this case, I assume that there are almost as many potential solutions as there are people studying arrhythmias. Winfree brings to the problem his set of geometric tools, so there is little surprise in his choice of the scroll ring as the target of his investigations. The reader must make his own judgement as to whether Winfree's solution is the right one.

The philosophy underlying the book is that topology has something to teach us about biology. Most of the examples deal with oscillatory processes in which there is a variable — call it phase — that can be

● Newly published by Harvard University Press is *Metabolic Arrest and the Control of Biological Time* by Peter W. Hochachka and Michael Guppy. The book examines the ability of some animals to regulate their own metabolism. Price is \$27.50; £21.95.

A NEW JOURNAL
just published from ISP:



the international journal
of
"AERIAL and SPACE
IMAGING, REMOTE
SENSING
and
INTEGRATED
GEOGRAPHICAL
SYSTEMS"

International advisory board now formed.

Regional Associate Editors:

S.A. Briggs, Farnborough, UK;
B. Dey, Washington D.C., USA;
B.C. Forster, Sydney, Australia;
P. Gudmandsen, Lyngby, Denmark;
M. Hallikainen, Helsinki, Finland
G. Ostrem, Oslo, Norway;
M.A. Pramanik, Dhaka, Bangladesh;
M. Sekioka, Kanagawa, Japan;
J.L. Star, Santa Barbara, USA;
K. Torlegard, Stockholm, Sweden;

In large format (A4) this new quarterly journal features original research and review papers, commercial and industrial applications, product news and supplement, advertising, institutional progress news, international news columns, special picture and interpretation features, etc. A wide spread of disciplines is covered, including planetary, solar and galactic imagery, upper air pollution monitoring, satellites for arms control, and crisis and disaster monitoring.

Send for a copy of the first issue from ISP. On subscription for Swed. kr. 800 (approx. US\$ 126) excl. postage, for libraries, institutions, companies, etc. On request reduced rate for personal use only



**Imaging Systems
Publications**

Karlavägen 18, 114 31 Stockholm, Sweden.
Tel. 08-205872. Telex 16830 SOS S

identified with a point of the circle. If the process is spread through space, one wants to associate each point with its phase; mathematically, one is defining a map whose image is the circle. A classical topological theorem states that there is no continuous mapping of a two-dimensional disk onto its boundary which leaves the points of its boundary fixed. With little additional mathematics, this theorem yields biological corollaries about singularities inside a spreading wave. Winfree is a master at finding and exploiting these corollaries in diverse biological settings; moreover he does so with a minimum of mathematics, relying instead on his rich geometric intuition.

As a mathematician, I am not altogether happy with this approach. The implicit assumption is that mathematical details and language are not needed to understand the basic argument Winfree puts forward. I question this. Three-dimensional topology is an intricate subject in which our geometric intuition is all too fallible. When mathematics is being used, is it here, to suggest the inevitability of observed phenomena, one must be careful to account for all of the qualifying clauses represented by the hypotheses of a theorem being applied. This is a difficult task to accomplish without the rigour that usually comes with mathematical jargon. Winfree's books are full of colourful and instructive figures, but I would have found this one more satisfying if there had been a few additional "boxes" to introduce and carry forward the mathematics more clearly.

Winfree makes a particular effort to pay attention to work that has been overlooked and can now be regarded as having been well ahead of its time. In view of this, I found it strange that the extensive work on the Belousov-Zhabotinsky reaction and its patterns by Nancy Kopell and Lou Howard throughout the 1970s receives no mention. There are also places at which the book is more dogmatic than appropriate in its accounts of controversial issues. One instance is the discussion of the shape of spiral waves in the Belousov-Zhabotinsky reagent (the latest experimental details can be found in Muller *et al.*'s article in *Physica* **24D**, 71-109; 1987).

Winfree's argument about the importance of topology in the organization of biological systems is hard to refute. He has pursued many examples in which we can see similar geometric patterns that appear to be related to topological singularities, and his book puts forward the proposition that sudden cardiac death is an instance of a topological disease. It is a case worth considering carefully. □

John Guckenheimer is in the Departments of Mathematics and Theoretical and Applied Mechanics, Cornell University, Ithaca, New York 14853, USA.

Beyond belief

David Miller

Forever Undecided: A Puzzle Guide to Gödel. By Raymond Smullyan. Knopf: 1987. Pp.247. \$17.95.

KNIGHTS always tell the truth, knaves never. You believe that, and that everyone is one or the other. Someone, knight or knave, declares "You will never believe that I am a knight". Then, if your complex of beliefs has some reasonable-sounding properties, you will remain forever undecided about whether the speaker is a knight. You can hardly believe that he is one, because that would involve believing what he says, which you know not to be true; but, unless you are quite confused about what you do believe, you cannot any more easily believe that he is a knave, for then you would disbelieve what he says. You will therefore stay undecided, which means that the speaker told the truth, and was a knight after all.

If for 'beliefs' we read 'theorems of a mathematical system' we may transform this little piece of whimsy into a version of Gödel's incompleteness theorem, perhaps the most significant result of twentieth century logic. In its improved form (due to Rosser) the theorem says that no consistent system of sufficient expressive power (the capability, in effect, of expressing a proposition that says of itself that it is not a theorem) can answer all questions that it can ask. In particular, it cannot settle the question of its own consistency. Gödel's theorem was later beautifully generalized by Martin Löb.

Through the medium of puzzles about knights and knaves — not to mention enigmatic professors of theology ("God exists if and only if you don't correctly believe that He does") and physicians who recommend faith healing ("If you believe that the cure will work, then it will work") — Smullyan escorts the reader through the subtleties of Gödel's theorem, and of

Löb's extension of it, to some deep results in modal logic, the logic of necessity and possibility. For all its light style, the book is not easy, and it will make heavy demands on a reader who has not already mastered elementary logic. There is much to be learnt from it, however, and the skilful presentation of Löb's work is especially illuminating.

Forever Undecided is Smullyan's fifth volume of recreational logic puzzles. It will be less of a success as a sweetener than its predecessors, I suspect, simply because there is too little sugar on the pill. Only a reader with much less self-awareness than the ideal reasoners featured in the puzzles will think that solving these puzzles is very different from proving theorems about formal systems. The various degrees of self-awareness that Smullyan investigates, for example, are worthy of consideration largely, if not only, because their analogues for mathematical systems are of interest. There is no independent rationale given for looking at them closely.

Reading the book is made more difficult by the absence of a glossary of technical terms, and even an index. Used with care, it could form the basis for a course on Gödel's and Löb's theorems, but there are some serious errors to beware of — the most astonishing is a hopelessly wayward truth table (p.42). It is worrying too that Smullyan insists on reading the material conditional " $p \supset q$ " as " p implies q " (p.39), the sin that Quine (quoted, but not named, on p.257) accused modal logic of being conceived in. An error of much less importance is that Smullyan disregards the possibility that someone might, by keeping silent, be both a knight and a knave.

There is, indeed, little in the book to please those interested in the history or the philosophy of logic. The discussion (pp.110-111) of Gödel's results on consistency, which fails to mention Hilbert, may serve as an example. □

David Miller is Senior Lecturer in the Department of Philosophy, University of Warwick, Coventry CV4 7AL, UK.



Overcrowding — microbial community on the surface of an estuarine sediment, dominated by the filamentous sulphur bacterium, *Beggiatoa*. From *Ecology of Protozoa: The Biology of Free-living Phagotrophic Protists* by Tom Fenchel, recently published in the Brock/Springer Series in Contemporary Bioscience, and distributed in USA, Canada and Mexico by Science Tech, Inc. Price is DM 94.00; \$39.00.

No business like show business

Michael Shortland

Interactive science centres are all the rage. But what — if anything — do children learn at them?

Two children, John and Mary, are at an interactive science centre. They are playing. After several moments having fun with whirling coloured discs, a ball of lightning and an assembly of distorting mirrors, they come upon an assortment of variously shaped blocks. They are invited to assemble them to form an arch. This they quickly do, and are encouraged to test the strength of the arch by bearing down on it. A slight wobble, then it stays firm. The children venture across the structure; it holds! Laughing, they hurry along to the next exhibit.

This is a familiar scene: a recognizable 'exploratory situation'. The first science centre, the Exploratorium, opened in San Francisco in 1969 and others were launched in North America, Australia and France during the 1970s. In Britain, Launch Pad, based in London's Science Museum, is doing brisk business; interactive centres have followed in Cardiff, Bristol and Liverpool, with others planned. For those who remember museums as hallowed places where visitors — especially children — were warned to hush and not touch, the new approach comes as a revelation. Gone are the fossils, speared insects and stuffed birds. Today's audience is encouraged to participate, discover and have fun. The talk is of 'plores', 'experiments' and 'games', as brigades of enthusiasts pull levers, press buttons, prod and bully exhibits into life. Cases and cabinets used to lie untouched, except by dust: the technician now anticipates heavy wear and tear, indeed encourages it. This is vandalism in a good cause, for as young visitors play they are learning about science.

Or are they? What do they actually learn? Are they receiving an educational experience in science, or about science, or, perhaps obliquely, about the process of learning itself? In the medley of promotional literature and enthusiastic press reports, it is by no means easy to decipher the nature of the benefits associated with science centres, but most advocates would claim that visitors are offered an experience in all three realms. They will be able

to grasp key scientific concepts and processes, sense the fascination of the scientific chase, and take home with them the message that education and entertainment *can* be combined.

The ingredients ('hands-on', 'lift-off') may be new, but the recipe is good old-fashioned fun. As John Bettelstone, Professor of Science Education and Honorary

equipment; children loved it. The trouble was that in the rush of enthusiasm, few bothered to query whether television *could* unite education and entertainment. Television became an end in itself rather than a means to an end. It produced a warped orientation towards education. Children came to love neither learning nor the subject being presented, but television

itself. The aim was to prepare children for entry into the exciting world of research: the achievement was to prepare them for entry into a fun-loving television culture. Far from improving as a result of television usage, learning in schools seems to have suffered^{2,3}. At best, youngsters were encouraged to enjoy science only if it resembled the make-believe world portrayed on the television screen.

There is surely a lesson here for all science centres. When education and entertainment are brought together under the same roof, education will be the loser. When John and Mary test the wooden arch, they are unlikely to delay to read the accompanying caption or wait for a gallery 'facilitator' to explain how the structure could bear their weight.

So what do children learn from an experience such as that? Probably very little from any literature pasted up around the display. Even if they take the time to read the caption, they may well find it long-winded and

over-technical. This is a common reaction, which perhaps says more about the difficulties of relating the child's lived experience to the adult's objective, sanctioned account of it than about the skills of the exhibitors themselves. The science writer John Gribbin recently took two 14-year-olds to the Exploratory in Bristol, established by Professor Richard Gregory. They loved the exhibits; "This was just as well," noted Gribbin, "since my two guinea pigs also said the captions on the different games were 'useless'. They were too long, boring and full of scientific jargon. But this didn't spoil the fun". Of course not, but one wonders whether it spoiled the education.

The real problem in this context is that

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The Bernoulli Blower. Reproduced courtesy of The Exploratory Hands-on Science Centre, Bristol.

Director of Cardiff's new Techniquet has said, "The main thing I want my visitors to say is 'I enjoyed that, it was fun'. If they don't I'm wasting my time". Like many other enthusiastic academics, Professor Bettelstone sees no problem in combining fact and fun: "We're in education, we're serious, but first and foremost we're part of the leisure and entertainment industry".

Attempts to conquer people's alienation from science are hardly new, nor are the particular approaches adopted by science centres. A decade ago, science educators were told to use television technology to present scientific information in an exciting, dramatic setting. Classrooms were equipped with the appropriate

no child wants to spoil the fun. Techniquest's brochure promises the visitor that " 'What', 'why', and 'how' will be words that often occur to you", but this is not the case. Children neither wait nor seem to welcome explanations; instead, understandably, they rush on to the next game. But what would happen if they did ask themselves such questions, say, "How do bridges stay up?". According to a leaflet put out by the Exploratory, the answer is "you can find out by building one". This will not do: putting blocks of wood together manifestly does not explain how a bridge or an arch holds firm. It may, to be sure, prompt an explanation; it may equally suggest a false one. Have the edges of the blocks of wood been coated with glue? Are there hidden devices (magnets perhaps?) buried in the blocks?

At interactive science centres children have fun participating in a series of 'experiments', but they learn little science and may acquire a good many misconceptions which at the very least fail to match those offered in the captions. No one can doubt the success of these centres in drawing visitors: "we're embarrassed by people's enthusiasm; we have crowd control problems", says Mike Williams, Launch Pad's Manager.

However, what can be questioned — and surely must be investigated — is whether any appropriate science is being acquired. To boast, as Launch Pad does, that its exhibits provide many opportunities for "experiencing fundamental scientific ideas" is to claim too much, and may even be absurd. After all, how does one experience an idea — gravity, say? By falling, walking, sitting?

The greatest potential for science centres lies in finding them a place in a broadly based educational initiative. In Britain, the integration of a mobile interactive fair within a BBC-sponsored Project Science in 1988 is particularly promising; this will provide an opportunity to re-orientate laypeople's views of science around a practical, problem-solving focus

"Science centres are an idea whose time has come in the late twentieth century. Everyone has welcomed them as nurseries of living thoughts, as opportunities to enhance scientific literacy . . . A worry nevertheless remains."

which has recognizable relevance. Two considerations should be paramount in this new travelling science centre: first, to offer exhibits or experiments which can suggest with the minimum of ambiguity certain scientific processes; second, to ensure that those processes are not simply seen as part of the world of science but as integral or related to everyday

phenomena.

Those familiar with established centres will recognize that these two objectives are not always achieved. Some of the science is far too complex to be rendered in exploratory form. In one popular exhibit, viewers are asked to peer into a mirror and wonder at the left-to-right, but not top-to-bottom, reversals that take place with reflection. Why does our head not appear where our feet are, when our arms and hands are transposed?

The question is wonderful, literally, but fails as a sensible problem to be explored by visitors. It has baffled such luminaries as Plato and Kant, caught out science writers like Martin Gardner, and has only recently been convincingly tackled by Professor Gregory himself. There is little hope that 11–14-year-olds will grasp the complexity of the dilemma; the chances are they will be left confused and disappointed.

If some explanations are too complicated, others are too simple. The 'Air Jet' exhibit is one of the most attractive. The Launch Pad programme explains: "Balance the beach ball on a vertical jet of air. Feel the Bernoulli force that keeps it in the jet even when it is tilted to one side". One may quibble with this description (the point is surely that the ball is not sitting *in* the jet of air; it appears to defy gravity because the pressure of the stationary air below it is higher than the pressure of the moving air above), but simplification goes one step further when it is claimed that this explains how aeroplanes fly. What it does is to illustrate (not explain) the effect that produces the pressure that lifts (not flies) aircraft.

The 'Air Jet' has the great merit of attempting to relate a piece of science (Bernoulli effect), technology (aeroplane) and a familiar phenomenon (flying). At a time when there is strong pressure from all quarters to make science education appropriate and to put it into context, such an exhibit shows how valuable a function science centres could perform. Nevertheless, far too many 'plores' purport to offer insights into science with precious little attempt to describe the material in its broader social and intellectual context. The contrast between this approach and the one taken by groups such as Professor Maurice Bazin's Living Science Space, based in Rio de Janeiro, is striking. Professor Bazin takes his troupe into the slums and shanty towns and involves residents in looking at the stars, examining pollution in the water supply, and analysing samples of faeces brought for on-the-spot diagnosis by workers and peasants¹. This is a far cry from the notion of science as entertainment, adventure and good clean fun.

Science is, of course, all of these things, but not these alone. It is in addition a

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Photograph reproduced by kind permission of the Science Museum, London.

serious, difficult, demanding enterprise which furthermore continually raises political and moral dilemmas. It is big business and frequently touches on sensitive vested interests, as Professor Bazin quickly discovered when his exploratory initiative took up issues connected with pollution as well as pulleys and pumps.

Science centres are an idea whose time has come in the late twentieth century. Everyone has welcomed them as nurseries of living thoughts, as opportunities to enhance scientific literacy, and as part of broader initiatives designed to promote public understanding of science. A worry nevertheless remains. What ideas and images of science (and, for that matter, of the scientist) are being cultivated here? Science as simply play, as innocent entertainment?

Science centres should strive to present an appealing yet balanced and realistic picture of their subject, and to nurture a critical, reflexive commitment to science and technology. There certainly is a tension between education and entertainment, but it is one that can be resolved by creating an atmosphere which is serious and enjoyable. What a pleasure it is to listen to the laughter and chatter, the thumps and splashes of children discovering the pleasures of our museums. What a greater pleasure it would be if we could be certain their brains were being taxed and tested at the same time. □

1. Schoon, N. *The Independent* (London), 25 May 1987.
2. Cohen, A. & Salomon, G. J. *Commun.* **29**, 156–163 (1979).
3. Jacoby, J., Hoyer, W.D. & Sheluga, D.A. *Miscomprehension of Televised Communications* (Education Foundation of the American Association of Advertising Agencies, New York, 1980).
4. Gribbin, J. *Guardian* (London & Manchester), 18 March 1987.
5. Bazin, M. *Sci. Literacy Pap.* **1**, 23–32 (1987).

Michael Shortland is in the Science Literacy Project, University of Oxford Department for External Studies, Rewley House, 1 Wellington Square, Oxford OX1 2JA, UK.

The strychnine-binding subunit of the glycine receptor shows homology with nicotinic acetylcholine receptors

Gabriele Grenningloh, Axel Rienitz, Bertram Schmitt, Christoph Methfessel*, Monika Zensen†§, Konrad Beyreuther†§, Eckart D. Gundelfinger & Heinrich Betz‡

Zentrum für Molekulare Biologie, Universität Heidelberg, Im Neuenheimer Feld 282, D-6900 Heidelberg, FRG

* Max-Planck-Institut für biophysikalische Chemie, Am Fassberg, D-3400 Göttingen, FRG

† Institut für Genetik, Universität zu Köln, Weyertal 121, D-5000 Köln 41, FRG

We have cloned and sequenced cDNAs of the strychnine-binding subunit of the rat glycine receptor, a neurotransmitter-gated chloride channel protein of the CNS. The deduced polypeptide shows significant structural and amino-acid sequence homology with nicotinic acetylcholine receptor proteins, indicating that there is a family of genes encoding neurotransmitter-gated ion channels.

NEUROTRANSMITTER-gated ion channels provide the molecular basis for rapid signal transmission at chemical synapses. So far, only one of these ion-channel proteins, the nicotinic acetylcholine receptor that mediates excitation of muscle and nervous tissue, has been analysed in detail. It is a hetero-oligomeric glycoprotein of homologous transmembrane polypeptides^{1,2}, assumed to have evolved by duplications of a common nicotinic-cholinergic ancestor gene^{3,4}. This paper shows that the ligand-binding subunit of an inhibitory central neurotransmitter receptor, the glycine receptor, also shares homology with the nicotinic acetylcholine receptor polypeptide family.

Synaptic inhibition in the central nervous system (CNS) of vertebrates is primarily mediated by two amino acids, γ -aminobutyric acid (GABA) and glycine⁵. Glycine is the major inhibitory neurotransmitter in spinal cord and brain stem, whereas GABA predominates in other parts of the brain. Binding of these amino acids to specific receptors increases the chloride conductance of the postsynaptic membrane and thus produces hyperpolarization (inhibition of neuronal firing)⁶. The convulsive alkaloid strychnine selectively antagonizes glycine-mediated, but not GABA-mediated inhibitory responses⁷. Glycine-displaceable ³H-strychnine binding has been widely used to investigate and localize the glycine receptor in different regions of the CNS⁸⁻¹⁰. From ³H-strychnine binding data, glycine receptor deficiencies have been implicated in the pathogenesis of spasticity^{11,12} and the loss of motor control associated with different human diseases, such as amyotrophic lateral sclerosis¹³ and Parkinsonism¹⁴.

The glycine receptor has been purified from spinal cord of different mammalian species by affinity chromatography on aminostrychnine-agarose and shown to be a large glycoprotein containing polypeptides of relative molecular masses (M_r) 48,000; 58,000 and 93,000 (48K, 58K and 93K)¹⁵⁻¹⁷. The 48K subunit can be covalently labelled with ³H-strychnine upon ultraviolet illumination¹⁸ and contains the antagonist binding site of the receptor¹⁹. Both the 48K and 58K polypeptides are glycosylated integral membrane components^{16,20}. Immunological crossreactivity and related peptide maps suggest that these polypeptides of the glycine receptor may be homologous²¹. The

Table 1 Inhibition of glycine receptor expression in oocytes by antisense oligonucleotides

Oligonucleotide injected	Glycine receptor channel activity responsive oocytes/oocytes tested
None	10/13
Nuc 1/20A	2/16
Nuc 1/20S	19/20
Nuc 35A	3/25

Preparation, injection, incubation and voltage-clamp analysis of *Xenopus laevis* oocytes were performed as described⁵³. Routinely, 30–50 nl of poly(A)⁺ RNA ($1 \mu\text{g} \mu\text{l}^{-1}$ in 100 mM NaCl, 10 mM Tris-HCl, pH 7.5) were injected per oocyte followed 1 h later by injection of 30–50 nl of oligonucleotide ($50 \text{ pmol} \mu\text{l}^{-1}$ in 100 mM NaCl, 10 mM Tris-HCl, pH 7.5). Oocytes were incubated for 5 days at 19 °C, and follicle cell layers were removed on the second day by collagenase treatment. Glycine receptor channel activity was determined in a two-electrode voltage clamp apparatus in response to bath application of 1 mM glycine. Holding potential was –90 mV. The amplitudes of the glycine-induced current varied between injected oocytes (range 25–90 nA). Oocytes exhibiting current signals <5 nA on glycine application were considered unresponsive. Oligonucleotide mixtures were synthesized by phosphoramidite chemistry⁵⁴ on an Applied Biosystems 380A DNA synthesizer with the following sequences: Nuc1/20A: 5'-ACNGCNCCTGTCYTCGTGCCA-3'; Nuc 1/20S: 5'-TGGCARGARCARGGNGCNGT-3'; Nuc 35A: 5'-CCATCAGCCACCTGCACG-GCNCCTGTCYTCGTGCCA-3' where N=all four nucleotides, R= purine nucleotides, and Y= pyrimidine nucleotides.

48K and 58K polypeptides have been proposed to form the ion-channel-containing 'core' of the receptor and to have evolved from a single chloride channel ancestor gene²¹⁻²³. The 93K polypeptide is a peripheral membrane protein localized at the cytoplasmic face of the postsynaptic glycine receptor complex^{20,24,25}. Here, we report the primary structure of the strychnine-binding 48K glycine receptor polypeptide as deduced from the complementary DNA sequence.

Cloning strategy

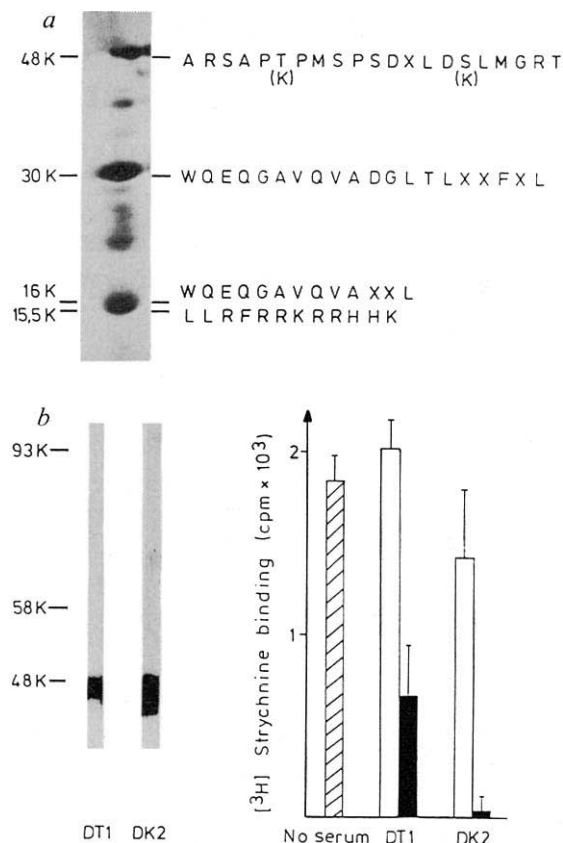
The strategy for cloning cDNAs for the 48K subunit of the rat glycine receptor was to screen cDNA libraries by hybridization with oligonucleotide mixtures synthesized on the basis of partial amino-acid sequence data. The 48K subunit was isolated from affinity-purified glycine receptor preparations¹⁵ by SDS-polyacrylamide gel electrophoresis. *Staphylococcus aureus* V8 protease cleaves the 48K subunit of the glycine receptor into a limited

‡ To whom correspondence should be addressed.

Present addresses: § Bayer AG, D-5090 Leverkusen, FRG (C.M.); Max-Planck-Institut für Psychiatrie, D-8033 Martinsried, FRG (M.Z.); Zentrum für Molekulare Biologie, Universität Heidelberg, D-6900 Heidelberg, FRG (K.B.).

Fig. 1 Partial amino-acid sequence analysis of the 48K glycine receptor polypeptide. *a*, *Staphylococcus aureus* V8 protease cleavage of 48K subunit and corresponding N-terminal polypeptide sequences identified by HPLC detection of phenylthiohydantoin (PTH) amino acids. X, unidentified amino acid position; residues in parentheses, additionally detected amino acids. *b*, Verification of the N-terminal sequence of the 48K polypeptide by anti-peptide antibodies. Left, immunoblot of affinity-purified glycine receptor. Right, immunoprecipitation of glycine receptor from detergent extracts of spinal cord membranes. The number of ^3H -strychnine binding sites in the extract was determined before (hatched bar) and after incubation with the indicated antiserum (solid bars), or preimmune serum (open bars).

Methods. Glycine receptor was purified from rat spinal cord membranes by affinity chromatography on aminostrychnine-agarose¹⁵, and the 48K polypeptide isolated by electrophoresis on a 10% SDS-polyacrylamide gel⁵⁵. For isolation of V8 protease fragments, excized gel pieces were squeezed on top of a 12–20% polyacrylamide gel and incubated in the gel matrix with *S. aureus* V8 protease⁵⁶. After electrophoresis, polypeptides were electroeluted from the gel and subjected to amino-acid sequence analysis on an Applied Biosystems 470A gas-phase sequencer equipped with a 120A PTH analyser⁵⁷. A peptide corresponding to the first 10 amino acids of the N-terminal region of the 48K polypeptide was synthesized with an additional tyrosine residue at the C terminus by Bachem (Switzerland). The peptide was coupled to thyroglobulin (serum DT1), or keyhole limpet haemocyanin (serum DK2), using bisdiazotated benzidine at a ratio of 1–1.5:1 (w/w) of peptide to carrier⁵⁸. After dialysis against water and 0.15 M NaCl, the conjugates were emulsified with Freund's complete adjuvant and used for the immunization of rabbits (1 mg per animal). Booster injections in incomplete adjuvant were given at intervals of 3–4 weeks. All sera tested contained anti-peptide antibodies as revealed by an enzyme-linked immunosorbent assay using the peptide as antigen (data not shown). For immunoblotting, receptor samples (1 μg per lane) were separated on a 10% SDS polyacrylamide gel, transferred to nitrocellulose and incubated sequentially with anti-peptide sera (diluted 100-fold) and anti-rabbit immunoglobulin G (IgG)²⁰. For immunoprecipitation, detergent extracts of rat spinal cord membranes⁴⁷ were incubated overnight at 4°C with protein A-Sepharose coated with antiserum, or control serum obtained from the same animals before immunization. ^3H -strychnine binding in the supernatant was determined as described⁴⁷.



number of peptides²¹. The 48K polypeptide and its major V8 protease fragments of 30K, 16K and 15.5K, respectively, were electroeluted from SDS-polyacrylamide gel slices and subjected to gas-phase microsequencing (Fig. 1*a*). The 30K and 16K peptides were found to have identical N-terminal sequences.

To confirm that the amino-acid sequences obtained were those of the 48K subunit, a decapeptide corresponding to its N-terminal sequence was synthesized, coupled to two different carrier proteins, and used to produce antisera. Two out of four sera obtained recognized the 48K polypeptide on immunoblots and precipitated ^3H -strychnine-binding sites from detergent extracts of rat spinal cord membranes (Fig. 1*b*).

Xenopus oocytes injected with brain poly(A)⁺ RNA express functional glycine receptor which is inserted into the oocyte membrane²⁶. We have exploited this expression system for the selection of oligonucleotide probes suitable for isolating cDNAs. A translation arrest assay similar to that described by Kawasaki²⁷ was used. Oocytes of developmental stage V were injected with poly(A)⁺ RNA isolated from spinal cord of 20-day-old rats. After 2–5 days of incubation, the injected oocytes developed a strong response to glycine. Under voltage clamp, bath application of glycine produced an inward current which could be blocked by strychnine. When an excess of a 20-mer oligonucleotide mixture corresponding to amino acids 1–7 of the N-terminal sequence of the 30K V8 protease fragment (Nuc 1/20A; see Table 1) was injected after application of the poly(A)⁺ RNA, the development of glycine sensitivity was almost completely prevented (Table 1). This blocking effect was specific as only the oligonucleotide mixture complementary to the messenger RNA, but not a 'sense' probe (Nuc 1/20S) or other oligonucleotide mixtures (not shown), inhibited expression of the receptor. Based on these data, we synthesized a further 35-mer mixed oligonucleotide corresponding to amino acids 1–12 of the 30K

V8 protease fragment (Nuc 35A). Up to base position 19, this probe covered all possible codon ranges; the subsequent bases were chosen according to codon usage and nucleotide frequency²⁸. This 35-mer anti-sense oligonucleotide probe also blocked glycine receptor expression in the oocyte system (Table 1).

Screening of 1.2×10^6 recombinant phages of a randomly primed $\lambda\text{gt}11$ cDNA library, prepared from rat spinal cord poly(A)⁺ RNA, with the radiolabelled oligonucleotides Nuc 1/20A and Nuc 35A revealed five independent recombinants which hybridized to both probes. One of these, clone GR1, contained a cDNA insert of 587 base pairs (bp) coding for the N-terminal amino-acid sequences of both the 30K and the 15.5K V8 protease-derived peptides (Fig. 2). Clone GR1 was used to isolate overlapping cDNA sequences from the $\lambda\text{gt}11$ library and from a $\lambda\text{gt}10$ library prepared from the same rat spinal cord poly(A)⁺ RNA using oligo(dT) as a primer. The clones analysed in establishing the entire coding DNA sequence of the mature 48K glycine receptor polypeptide are shown in Fig. 2*a*.

Sequence and tissue location of mRNA

The cDNA sequence of 1,384 nucleotides that encode the 48K polypeptide of the rat glycine receptor (Fig. 2*b*) was determined using clones GR1 (nucleotides 449–1,035), GR2 (nucleotides 302–952), GR1-6 (nucleotides 379–1,384) and GR1-6-1 (nucleotides 1–709). Except for a single C to T transition in nucleotide position 561, all four clones are identical in their overlapping regions. The partial amino-acid sequences determined (Fig. 1*a*) are encoded by the cDNA in the same reading frame (amino-acid residues 1–21, 170–189 and 314–325). Two incorrect assignments were observed at amino-acid positions 6 and 16 which probably arose from a low recovery of PTH-derivates in the respective Edman degradation cycles. Using the reading frame defined by

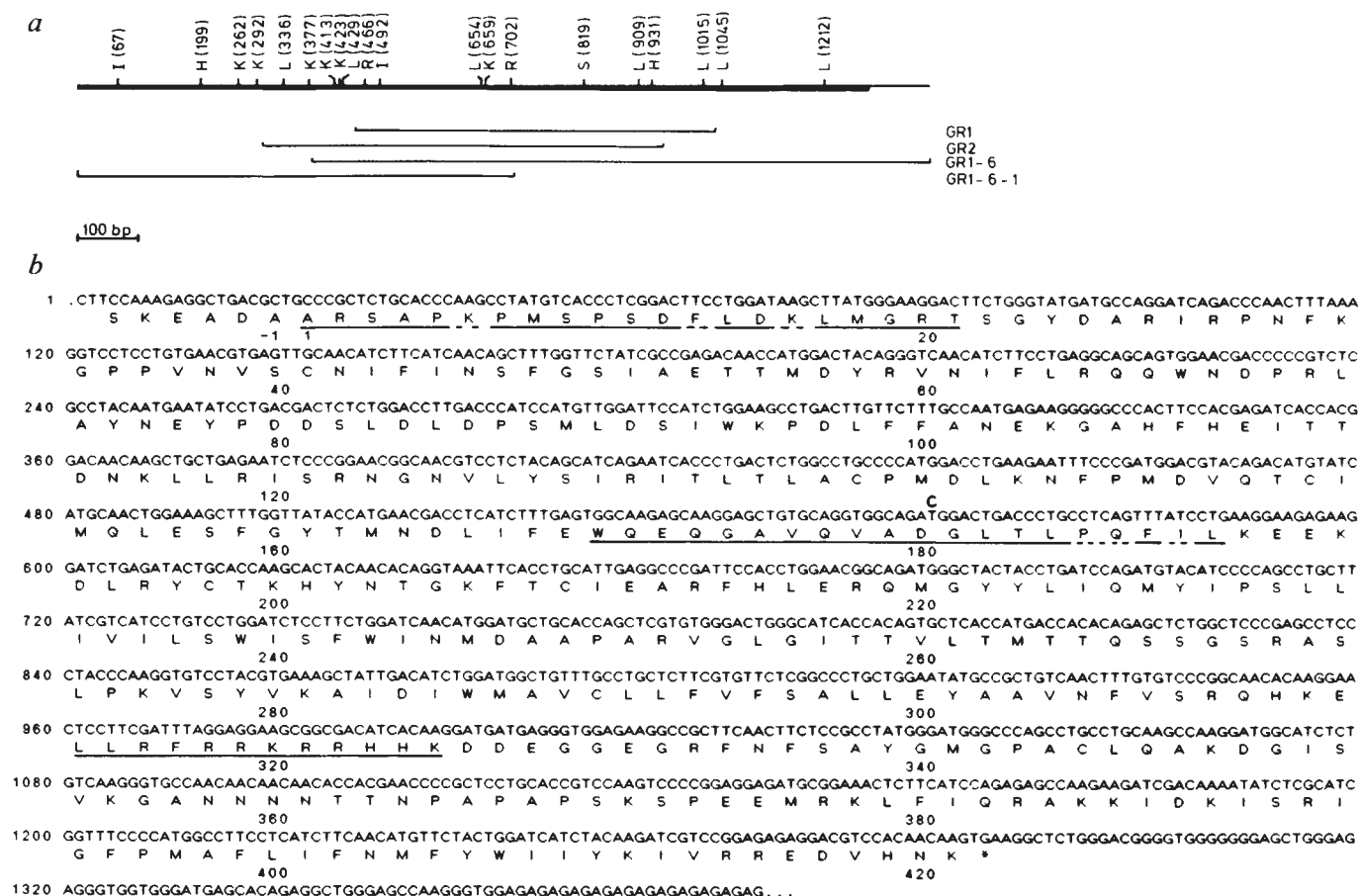


Fig. 2 *a*, Restriction map of the rat glycine receptor 48K subunit cDNA clones. Only the relevant restriction endonuclease sites are shown (H, *Hind*II; I, *Hind*III; K, *Hin*fI; L, *Hae*III; R, *Rsa*I; S, *Ssr*I). Solid box, protein coding region. *b*, Nucleotide and deduced amino-acid sequence of the 48K subunit cDNAs. Nucleotides are numbered in the 5' to 3' direction at left, and amino acids below the sequences. Numbering of amino acids starts at the first residue of the mature N-terminal sequence of the 48K polypeptide. Solid lines, amino-acid sequences verified by partial peptide sequence analysis; broken lines, uncertain or not identified residues.

Methods. Total cellular RNA was isolated from spinal cord tissue of three-week-old rats by the guanidinium isothiocyanate-CsCl method⁵⁹. Poly(A)⁺ RNA was purified by oligo(dT)-cellulose chromatography⁶⁰. Single-stranded cDNA was synthesized using oligo(dT) or a random hexamer as primers⁶¹. Second-strand synthesis was carried out using *Escherichia coli* DNA polymerase I (Klenow fragment) followed by S₁ nuclease treatment⁶¹. The ends of the cDNA were blunted with *E. coli* DNA polymerase I (Klenow fragment) and, following methylation of the *Eco*RI sites, *Eco*RI linkers were added, and the cDNA was ligated into the vector λ gt10 (ref. 62). The construction of the randomly primed λ gt11 library will be described in detail elsewhere. About 1.2×10^6 recombinant phages from the amplified λ gt11 library were screened with the ³²P-labelled oligonucleotides Nuc 1/20A and Nuc 35A (Table 1). Hybridization with Nuc 1/20-A was carried out at 42 °C in 3 M TMACl (tetramethylammonium chloride), 50 mM Tris, pH 8.0, 2 mM EDTA, 10× Denhardt's solution and 200 µg ml⁻¹ transfer RNA. Filters were washed in 3 M TMACl, 50 mM Tris, pH 8.0, 2 mM EDTA and 0.1% SDS at 52 °C. Hybridization with Nuc 35A was performed as described²⁸. For rescreening of the λ gt10 and λ gt11 libraries with clone GR1 and the 5' *Hind*III fragment of the insert of clone GR1-6 (nucleotides 379–492), respectively, the gel-purified DNA fragments were labelled with ³²P-dCTP by nick-translation⁶¹ and hybridized at 65 °C (ref. 63). The cDNA inserts of hybridization-positive phages were subcloned into pUC18, or M13mp18 and mp19 vectors⁶⁴, for DNA sequence analysis by the dideoxy chain termination method⁶⁵. Both strands of the cDNA clones were sequenced.

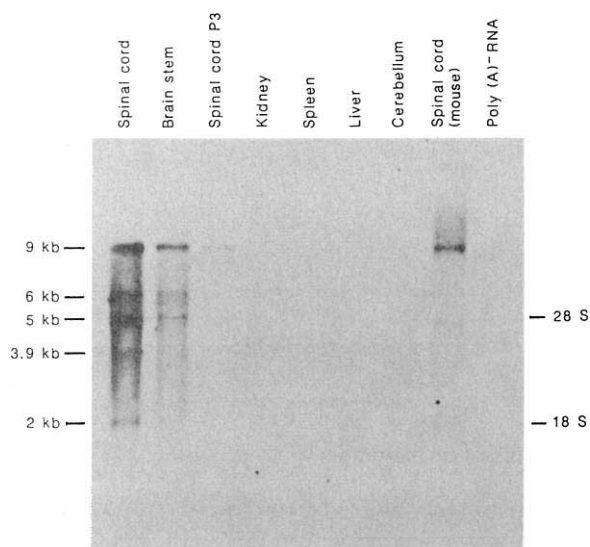
the peptide sequences, the complete amino-acid sequence of the 48K glycine receptor subunit was deduced (Fig. 2*b*). The N-terminal amino-acid alanine of the mature 48K polypeptide was assigned to nucleotides 21–23 of the cDNA. An in frame TGA translation termination codon occurs at nucleotides 1,284–1,286. The mature 48K subunit of the rat glycine receptor thus consists of 421 amino-acid residues and has a calculated M_r of 48,383.

The cDNA sequence given here lacks the 5' untranslated region and most of the sequence encoding the presumptive signal peptide of the 48K subunit precursor. It also contains only a minor portion of the 3' non-coding region of the corresponding mRNA: 101 nucleotides which, between nucleotides 1,361 and 1,384, contain 12 simple AG repeats. By Northern blot analysis, two major transcripts of ~9.0 and ~5.0 kilobases (kb) were detected in poly(A)⁺ RNA preparations of rat spinal cord and brain stem, but not of kidney, liver, spleen or cerebellum (Fig.

3). The cerebellum contains relatively low levels of glycine receptor⁸⁻¹⁰. In poly(A)⁺ RNA from spinal cord of 20-day-old rats, additional hybridizing bands of 6 kb, 3.9 kb and 2 kb were observed. Spinal cord poly(A)⁺ RNA from 3-day-old rats contained small amounts of 48K subunit mRNA. At this stage of development, there are only a few ³H-strychnine binding sites in spinal cord²⁹. Poly(A)⁺ RNA of mouse spinal cord gave hybridization signals at 9 kb, 4.9 kb and 2 kb, respectively.

Structure of the deduced protein

The predicted amino-acid sequence of the mature 48K polypeptide is preceded by residues which probably belong to the signal peptide of its precursor. Although not typical of eukaryotic signal sequences, they provide a C-terminal alanine suitable for cleavage by signal peptidase³⁰. Our data do not, however, exclude the possibility that there is an N-terminally extended



pro-form of the polypeptide. As expected for a glycoprotein, a potential N-glycosylation site is found at amino-acid position 38 of the 48K subunit^{16,20} (Fig. 2b). Analysis for regional hydrophathy³¹ revealed four hydrophobic segments long enough to form transmembrane α -helices between amino acids 220 and 246, 253 and 270, 285 and 308, and 393 and 410, respectively (compare underlined sequences in Fig. 4a). This arrangement of possible transmembrane regions is similar to that of nicotinic acetylcholine receptor subunits¹⁻⁴, so both types of receptor polypeptides may have the same transmembrane topology. Significant amino-acid sequence homology exists between the 48K glycine receptor subunit and different members^{1-4,32-38} of the nicotinic acetylcholine receptor polypeptide family. An alignment of the 48K subunit sequence with the nicotinic acetylcholine receptor protein ARD from *Drosophila melanogaster*³⁶, which exhibits the highest homology to the glycine receptor polypeptide amongst neuronal cholinergic receptors, is shown in Fig. 4a. Of all the amino-acid residues, 21% are identical in both proteins; another 22% result from conservative exchanges. Homology is high in the presumptive extracellular region and transmembrane segments M1 to M3. Here, 18-24% identical amino-acid positions are observed between residues 11 and 309 of the 48K glycine receptor polypeptide and the respective portion of nicotinic acetylcholine receptor α , β , γ , δ and ϵ subunits of *Drosophila*³⁶, *Torpedo*^{3,37}, chicken³⁴, mouse³², rat^{4,35,38}, cow^{33,37} and human³⁷ (not shown). Particularly striking is a region around two cysteines found at precisely conserved positions (amino acids 140 and 154 in Fig. 4a; the 'conserved Cys-Cys domain') of the presumptive extracellular portions of the 48K and ARD polypeptides. Ten out of 17 amino acids in this region are identical in both proteins. This domain is highly conserved between nicotinic acetylcholine receptor proteins^{3,4,32-38}. Its two cysteines are thought to form a disulphide bridge essential for tertiary structure formation of the acetylcholine receptor^{3,33,35,37,39}.

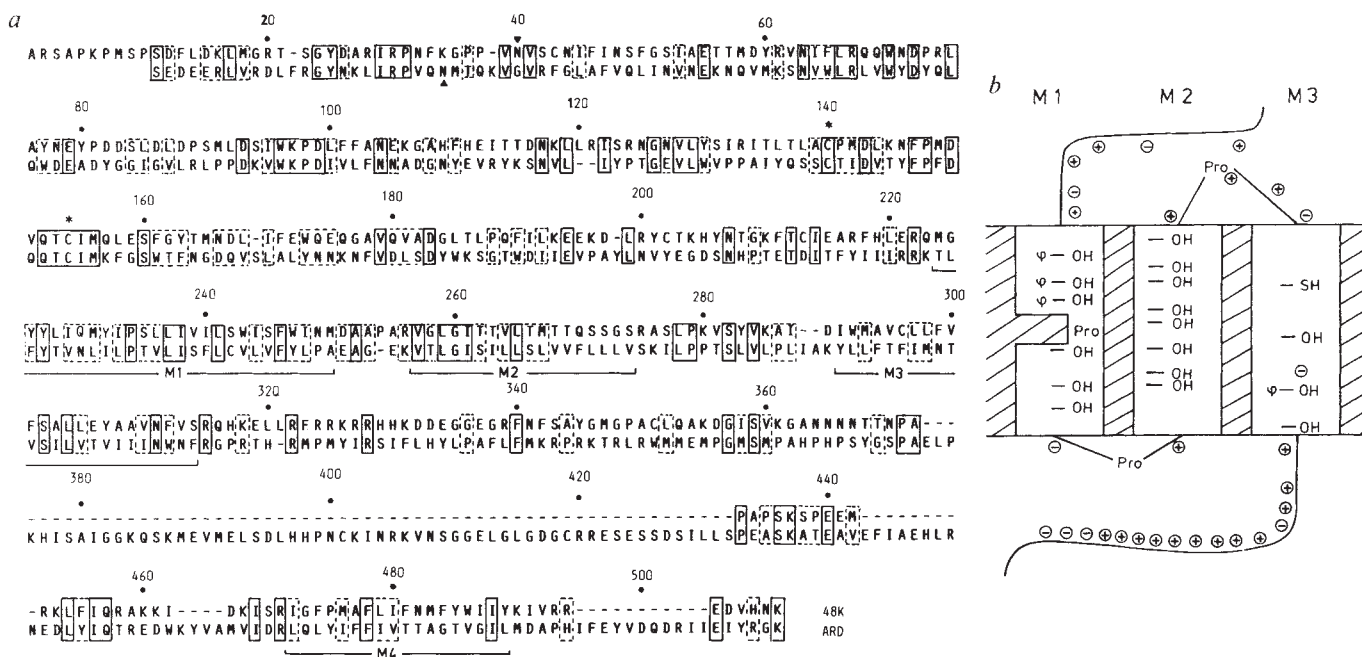
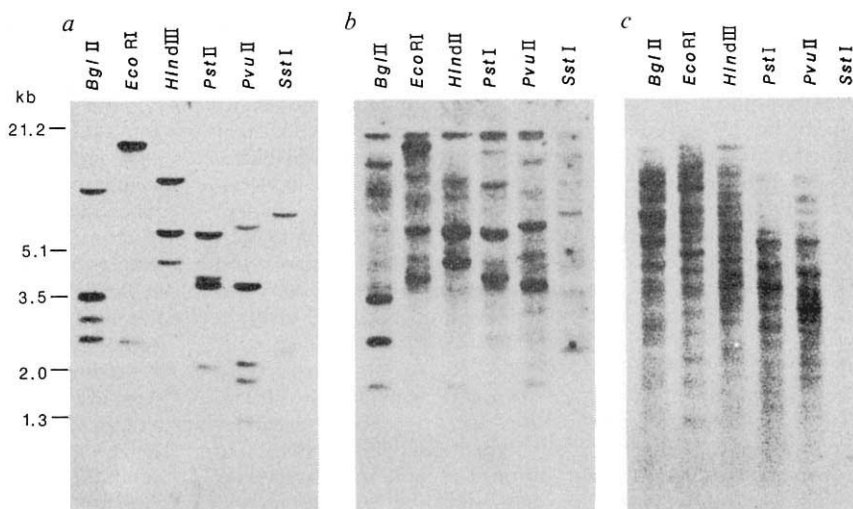


Fig. 5 Southern analysis of genomic DNA with 48K glycine receptor subunit cDNA probes. Rat genomic DNA was digested with the indicated restriction enzymes⁶¹. After separation of the products in a 0.8% agarose gel and transfer to a nylon membrane (Zetabind, AMF), the blot was hybridized sequentially to the following probes which were nick-translated to specific activities of 1×10^8 c.p.m. μg^{-1} . *a*, cDNA insert of clone GR2 under conditions of high stringency⁶³ in buffer containing 50% formamide; *b*, *HinfI*-*EcoRI* fragment of clone GR2 (nucleotides 659-952) corresponding to the transmembrane regions M1-M3 under conditions of low stringency⁶³ using 43% formamide in the hybridization buffer. *c*, *HaeIII*-*HindIII* fragment (nucleotides 429-492) of clone GR2 corresponding to the conserved Cys-Cys domain using low-stringency hybridization conditions as in *b*. DNA of phage λ digested with *EcoRI* and *HindIII* was used as a size marker.



The hydrophobic segments of the 48K glycine receptor subunit, in particular M1 to M3 (Fig. 4*b*), may be part of the chloride channel of the glycine receptor. Segments M1 and particularly M2 contain many uncharged polar amino-acid residues (Fig. 4*b*). These segments thus may form part of the hydrophilic inner wall of the glycine receptor channel. Interestingly, the transmembrane segment M2 of different nicotinic acetylcholine receptor subunits has been shown to be involved in cation transport⁴⁰ and channel blocker binding^{41,42}. Therefore segment M2 may be crucial for ion translocation in different types of neurotransmitter-gated ion channels. A proline residue conserved in segment M1 of nicotinic acetylcholine receptor subunits^{3,4,32-38} is also found in the 48K glycine receptor polypeptide (Fig. 4*b*). Proline residues interrupting transmembrane-spanning α -helices may provide the structural flexibility required for reversible conformational transitions of intramembrane regions of transport proteins through *cis-trans* isomerization of X-Pro peptide bonds⁴³ (where X is any amino acid). Segment M3 contains a potentially charged amino-acid residue (Fig. 4*b*). Segment M4 mainly consists of hydrophobic amino acids and thus is unlikely to contribute to the ion channel proper.

The bend portions connecting segments M1 and M2, and M2 and M3, respectively, both contain a proline residue (Fig. 4*b*). The extracellular sequence preceding segment M1, the connecting region between segments M2 and M3, and the intracellular sequence following M3 have several positively charged amino-acid residues. A stretch of eight uninterrupted basic residues followed by three negatively charged amino acids is found close to the C-terminal end of segment M3. Patch clamp data indicate that there are two sequentially occupied anion binding sites in the chloride channel of both the glycine and the GABA receptor⁴⁴. The positively charged residues surrounding segments M1 to M3 may serve as sites of this type at the presumptive inner and outer mouths of the ion channel and provide its ion selectivity⁴⁴. In addition they may contribute to its voltage-dependent conductance properties^{44,45} and regulation by intracellular chloride⁴⁶.

For the nicotinic acetylcholine receptor, it has been proposed that an amphipathic α -helix preceding segment M4 forms the inner wall of its ion channel². The protein sequence of the 48K glycine receptor polypeptide does not contain a corresponding region suitable for the formation of an amphipathic α -helix.

The agonist binding site of the glycine receptor has not been identified as no suitable high-affinity ligands are available. However, based on the following considerations we tentatively assign it to amino-acid residues 190-196 of the 48K subunit.

First, strychnine and glycine interact at the glycine receptor in an apparently competitive fashion^{8,17,47}. Protein modification experiments indicate that their binding sites are closely associated, but not identical⁴⁸. Second, the ultraviolet-radiation induced labelling of the 48K polypeptide with ³H-strychnine has been shown to require energy transfer from aromatic side chains¹⁹. Third, chemical modification of tyrosine residues inhibits ³H-strychnine binding to membrane-bound glycine receptor, whereas amino-group modification abolishes its displacement by glycine⁴⁸. Fourth, protease treatment cleaves ≥ 11 K of the extracellular portion of the 48K polypeptide without removing photo-incorporated ³H-strychnine from the membrane-bound receptor¹⁹. Therefore, the ligand binding site cannot be within the first 100 amino acids of the 48K subunit. Between amino-acid 100 and segment M1, tyrosines are found at positions 128, 161, 197 and 202, respectively. Residues 197 and 202 are surrounded by charged amino acids which may mediate binding of the alkaloid strychnine. Three acidic and three basic residues (amino acids 190-194 and 196) precede tyrosine 197. These may form a secondary structure which binds one, or even more⁴⁸, glycine molecules. The ligand-binding region of the glycine receptor thus may be located between amino acids 190 and 202 of the 48K subunit. This position corresponds to that of the acetylcholine binding region on the α -subunit of cholinergic receptor proteins^{1,2,39,49}. Also, both the acetylcholine binding site of the nicotinic receptors^{3,4,35,37,49,50} and the ligand-binding site of the glycine receptor proposed here contain two adjacent, or closely spaced, cysteine residues. In the case of the 48K polypeptide these cysteines (residues 198 and 209) may form a disulphide-bonded loop which exposes amino-acid side chains involved in antagonist binding.

Related genes

A Southern blot of rat genomic DNA digested with various restriction enzymes was hybridized with radiolabelled cDNA of clone GR2, which includes the conserved Cys-Cys domain, the presumptive ligand-binding region and transmembrane regions M1 to M3. Under conditions of high stringency, few hybridizing bands were observed with the differently cut rat genomic DNA (Fig. 5*a*). The gene of the 48K subunit thus presumably is present only once in the haploid rat genome. Southern blots of the same DNA preparations were also probed under hybridization conditions of low stringency, using cDNA probes encompassing either the transmembrane regions M1 to M3 or the conserved Cys-Cys domain. With the transmembrane probe, some additional hybridizing bands were observed which

were not seen with clone GR2 under high-stringency conditions (Fig. 5b). These may correspond to other putative chloride channel genes. Low-stringency hybridization with the probe corresponding to the conserved Cys-Cys domain revealed many hybridizing DNA fragments (Fig. 5c). This result is consistent with the idea that there is a large family of neurotransmitter receptor proteins in the rat which includes glycine and acetylcholine and potentially also GABA and glutamate receptor polypeptides. The combined use of DNA probes corresponding to both the transmembrane regions of ligand-gated ion channels and the conserved Cys-Cys domain therefore may allow identification of new neurotransmitter receptor genes and cDNAs.

Conclusions

The cDNA sequence data presented here show that the strychnine-binding 48K subunit of the glycine receptor resembles nicotinic acetylcholine receptor polypeptides in both amino-acid sequence and structural organization. The subunits of both types of neurotransmitter receptors presumably evolved from a common ancestor protein. Important characteristics of primary structure are shared by the 48K subunit and the nicotinic acetylcholine receptor proteins. First, they have four hydrophobic segments at similar positions, and thus most probably the same transmembrane topology. Second, they possess a highly conserved Cys-Cys domain in their extracellular region. Its function is presently unknown, but may relate to the structure, cellular localization, or interaction with extracellular components, of these plasma membrane proteins. Third, the relative distribution of both charged and non-charged polar amino acids in and around the putative ion-channel-forming transmembrane regions of these receptor polypeptides is similar. Furthermore, the position of the presumptive ligand-binding sites on the 48K glycine receptor subunit and the acetylcholine receptor α -subunits is not only conserved, but two additional closely spaced

cysteines are found here in both types of receptor polypeptides. Apparently, the common structural features of these receptor subunits reflect their close functional similarity. Both the glycine and the nicotinic acetylcholine receptor form neurotransmitter-gated ion channels. The different nature of their agonists and permeating ions obviously requires only minor modifications of a common protein architectural theme. Our hybridization analysis of rat genomic DNA in addition suggests that the subunits of other ligand-gated ion channels also may be members of the same protein family. All the different neurotransmitter receptors in the nervous system thus may belong to only two large gene families: the ion channel type described here and the rhodopsin-like receptors^{51,52} which couple to G proteins.

During the preparation of this manuscript, the groups of E. Barnard and P. Seeburg (personal communication) reported the cloning of GABA receptor polypeptides at a recent conference. Their data indicate that the subunits of the GABA receptor exhibit structural homology to the 48K glycine receptor polypeptide and nicotinic acetylcholine receptor proteins.

We thank B. Sakmann for support with the oocyte experiments, A. Bach for help with the preparation of cDNA libraries; H. Krischke, C. Udri and C. Schröder for technical assistance, C.-M. Becker, W. B. Huttner and C. H. Schaller for a critical reading of the manuscript; and J. Rami and I. Nonnenmacher for help with its preparation. This work was supported by the Bundesministerium für Forschung und Technologie, the Deutsche Forschungsgemeinschaft, Forschungsschwerpunkt 18 des Landes Baden-Württemberg, and the Fonds der Chemischen Industrie. Provision of an Applied Biosystems DNA synthesizer by the Fonds der Chemischen Industrie is gratefully acknowledged.

Note added in proof: The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00276.

Received 21 April; accepted 29 May 1987.

- Popot, J. L. & Changeux, J. P. *Physiol. Rev.* **64**, 1162-1239 (1984).
- Stroud, R. M. & Finer-Moore, J. A. *Rev. Cell Biol.* **1**, 317-351 (1985).
- Noda, M. *et al. Nature* **305**, 818-823 (1983).
- Boulter, J. *et al. Nature* **319**, 368-374 (1986).
- Krnjević, K. *Physiol. Rev.* **54**, 418-540 (1974).
- Coombs, J. S., Eccles, J. C. & Fatt, P. J. *Physiol., Lond.* **130**, 326-373 (1955).
- Aprison, M. H. & Werman, R. *Life Sci.* **4**, 2075-2083 (1965).
- Young, A. B. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2832-2836 (1973).
- Zarbin, M. A., Wamsley, J. K. & Kuhar, M. J. *J. Neurosci.* **1**, 532-547 (1981).
- Probst, A., Cortes, R. & Palacios, J. M. *Neuroscience* **17**, 11-35 (1986).
- Hall, P. V., Smith, J. E., Lane, J., Mote, J. & Campbell, R. *Neurology* **29**, 262-267 (1979).
- White, W. F. & Heller, A. H. *Nature* **298**, 655-657 (1982).
- Hayashi, H., Suga, M., Satake, M. & Tsubaki, T. *Ann. Neurol.* **9**, 292-294 (1981).
- Lloyd, K. G., De Montis, G., Broekkamp, C. L., Thuret, F. & Worms, P. *Adv. biochem. Psychopharmac.* **37**, 137-148 (1983).
- Pfeiffer, F., Graham, D. & Betz, H. *J. biol. Chem.* **257**, 9389-9393 (1982).
- Graham, D., Pfeiffer, F. & Betz, H. *Biochemistry* **24**, 990-994 (1985).
- Becker, C.-M., Hermans-Borgmeyer, I., Schmitt, B. & Betz, H. *J. Neurosci.* **6**, 1358-1364 (1986).
- Graham, D., Pfeiffer, F. & Betz, H. *Biochem. biophys. Res. Commun.* **102**, 1330-1335 (1981).
- Graham, D., Pfeiffer, F. & Betz, H. *Eur. J. Biochem.* **131**, 519-525 (1983).
- Schmitt, B., Knaus, P., Becker, C.-M. & Betz, H. *Biochemistry* **26**, 805-811 (1987).
- Pfeiffer, F., Simler, R., Grenningloh, G. & Betz, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7224-7227 (1984).
- Betz, H. *Naturwissenschaften* **71**, 363-368 (1984).
- Betz, H. *Trends Neurosci.* **10**, 113-117 (1987).
- Triller, A., Cluzaud, F., Pfeiffer, F., Betz, H. & Korn, H. *J. Cell Biol.* **101**, 683-688 (1985).
- Altschuler, R. A., Betz, H., Parakkal, M. H., Reeks, K. A. & Wenthold, R. J. *Brain Res.* **369**, 316-320 (1986).
- Simikawa, K., Parker, I. & Miledi, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7994-7998 (1984).
- Kawasaki, E. S. *Nucleic Acids Res.* **13**, 4991-5004 (1985).
- Lathe, R. J. *J. molec. Biol.* **183**, 1-12 (1985).
- Benavides, J., Lopez-Lahoya, J., Valdivieso, F. & Ugarte, M. *J. Neurochem.* **37**, 315-320 (1981).
- Hortsch, M. & Meyer, D. I. *Int. Rev. Cytol.* **102**, 215-245 (1986).
- Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824-3828 (1981).
- La Polla, R. J., Mixer-Mayne, K. N. & Davidson, N. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7970-7974 (1984).
- Kubo, T. *et al. Eur. J. Biochem.* **149**, 5-13 (1985).
- Nef, P., Mauron, A., Stadler, R., Alliod, C. & Ballivet, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7975-7979 (1984).
- Boulter, J. *et al. J. Neurosci.* **5**, 2545-2552 (1985).
- Hermans-Borgmeyer, I. *et al. EMBO J.* **5**, 1503-1508 (1986).
- Numa, S. *Biochem. Soc. Symp.* **52**, 119-143 (1986).
- Goldman, D. *et al. Cell* **48**, 965-973 (1987).
- Mishina, M. *et al. Nature* **313**, 364-369 (1985).
- Imoto, K. *et al. Nature* **324**, 670-674 (1986).
- Giraudat, J., Dennis, M., Heidmann, T., Chang, J.-Y. & Changeux, J.-P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2719-2723 (1986).
- Oberthür, W. *et al. EMBO J.* **5**, 1815-1819 (1986).
- Brandl, C. J. & Deber, C. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 917-921 (1986).
- Bormann, J., Hamill, O. P. & Sakmann, B. *J. Physiol., Lond.* **385**, 243-286 (1987).
- Faber, D. S. & Korn, H. *J. Neurosci.* **7**, 807-811 (1987).
- Gold, M. R. & Martin, A. R. *Nature* **299**, 828-830 (1982).
- Pfeiffer, F. & Betz, H. *Brain Res.* **226**, 273-279 (1981).
- Young, A. B. & Snyder, S. H. *Molec. Pharmacol.* **10**, 790-809 (1974).
- Kao, P. N. *et al. J. biol. Chem.* **259**, 11662-11665 (1984).
- Noda, M. *et al. Nature* **299**, 793-797 (1982).
- Dixon, R. A. F. *et al. Nature* **321**, 75-79 (1986).
- Kubo, T. *et al. Nature* **323**, 411-416 (1986).
- Methfessel, C. *et al. Pflügers Arch. ges. Physiol.* (in the press).
- Sinha, N. D., Biernat, J., McManns, J. & Kosters, H. *Nucleic Acids Res.* **12**, 4539-4557 (1984).
- Laemmli, U.K. *Nature* **227**, 680-685 (1970).
- Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U.K. *J. biol. Chem.* **252**, 1102-1106 (1977).
- Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* **256**, 7990-7997 (1981).
- Bassiri, R. M. & Utiger, R. D. *Endocrinology* **90**, 722-727 (1971).
- Kaplan, B. B., Bernstein, S. L. & Gioio, A. *Biochem. J.* **183**, 181-184 (1979).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbour Laboratory, New York, 1982).
- Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning, a Practical Approach* Vol. 1 (ed. Glover, D. M.) 49-78 (IRL, Oxford, 1984).
- Mason, P. J. & Williams, J. G. *Nucleic Acid Hybridization, a Practical Approach* (eds Hames, D. & Higgins, J.) 113-137 (IRL, Oxford, 1985).
- Norlander, J., Kempe, T. & Messing, J. *Gene* **26**, 101-106 (1983).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
- Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. in *Atlas of Protein Sequence and Structure* Vol. 5, suppl. 3 (ed. Dayhoff, M. O.) 345-352 (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).
- Hubbard, S. C. & Ivatt, R. J. A. *Rev. Biochem.* **50**, 555-583 (1981).

Sequence and functional expression of the GABA_A receptor shows a ligand-gated receptor super-family

Peter R. Schofield^{*§}, Mark G. Darlison[†], Norihisa Fujita[†], David R. Burt^{†§},
F. Anne Stephenson[†], Henry Rodriguez^{*}, Lucy M. Rhee^{*}, J. Ramachandran^{*},
Vincenzina Reale[†], Thora A. Glencorse[†], Peter H. Seeburg^{*§} & Eric A. Barnard^{†‡}

^{*} Genentech, Inc., Department of Developmental Biology, 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

[†] MRC Molecular Neurobiology Unit, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

*Amino-acid sequences derived from complementary DNAs encoding the α - and β -subunits of the GABA/benzodiazepine receptor from bovine brain show homology with other ligand-gated receptor subunits, suggesting that there is a super-family of ion-channel-containing receptors. Co-expression of the in vitro-generated α -subunit and β -subunit RNAs in *Xenopus* oocytes produces a functional receptor and ion channel with the pharmacological properties characteristic of the GABA_A receptor.*

γ -AMINOBUTYRIC acid (GABA), the major inhibitory neurotransmitter in the vertebrate brain, mediates neuronal inhibition by binding to the GABA/benzodiazepine (GABA_A) receptor and opening an integral chloride channel. The structure of this receptor is of particular interest because it contains a variety of binding sites for pharmaceutically significant drugs which interact allosterically with the GABA agonist site or the receptor channel. These include anxiolytic (benzodiazepines), anti-convulsant (barbiturates), anxiogenic (β -carbolines) and convulsant (picrotoxin) agents (reviewed in refs 1 and 2). The receptor complex has been affinity-purified from bovine cerebral cortex with all of those binding sites intact^{3,4}. The subunit structure of the complex has been shown^{5,6} to be $\alpha_2\beta_2$. Photo-affinity labelling experiments on the pure receptor have demonstrated that the α -subunit (apparent relative molecular mass 53,000; M_r 53K) carries the benzodiazepine binding site and the β -subunit (apparent M_r 57K) the GABA binding site^{5,6}. Limited structural similarity is suggested by the recent finding⁶ that only one of about twenty^{6,7} monoclonal antibodies raised so far against this receptor recognizes both the α - and the β -chain.

We have used this purified receptor to generate peptide sequences, allowing the isolation of cloned cDNAs encoding both subunits. We here report the deduced amino-acid sequences of two GABA_A-receptor subunits, which we find to have homologies in both primary sequence and predicted transmembrane topology. Comparison of these sequences with those of other neurotransmitter receptors identifies structural features which appear characteristic of the chemically-gated ion channels. Further, we demonstrate by functional expression in *Xenopus* oocytes that the cloned subunit cDNAs encode the entire GABA_A receptor with its full pharmacological profile.

cDNA and protein sequences

The GABA_A receptor was purified to homogeneity from bovine cerebral cortex by benzodiazepine-affinity chromatography, essentially as described previously⁴. The purified preparation bound [³H]muscimol with high affinity and comprised only α - and β -subunits when analysed by SDS-polyacrylamide gel electrophoresis under reducing conditions, as previously found^{3,4}. Gas-phase micro-sequencing of the entire protein failed to yield detectable signals, presumably due to blocked N-termini. But

cleavage of the whole receptor by cyanogen bromide and HPLC separation of the peptides resulted in several amino-acid sequences, of which three (CN8.9, CN10.8 and CN10.9) were used for the design of synthetic oligodeoxyribonucleotide probes (see Fig. 1 legend). In addition, the α -subunit protein was isolated by electroelution from SDS-polyacrylamide gels and, after tryptic cleavage, a further α -specific peptide sequence (T33) was obtained.

Bovine brain and calf cerebral cortex cDNA libraries, constructed⁸ in phage λ gt10, were screened with the synthetic probes. Hybridizing clones were analysed by DNA sequencing and found to be of two types. One specified an open reading frame which contained three of the chemically-determined peptide sequences, including the one specific for the α -subunit (CN8.9, CN10.8 and T33; see Fig. 1). The other type of clone did not contain the α -specific sequence (T33) but did contain the coding sequence for peptide CN10.9 (not found in the α -subunit cDNA) and was therefore designated as encoding the β -subunit. The nucleotide and deduced amino-acid sequences of both subunits are shown in Fig. 1.

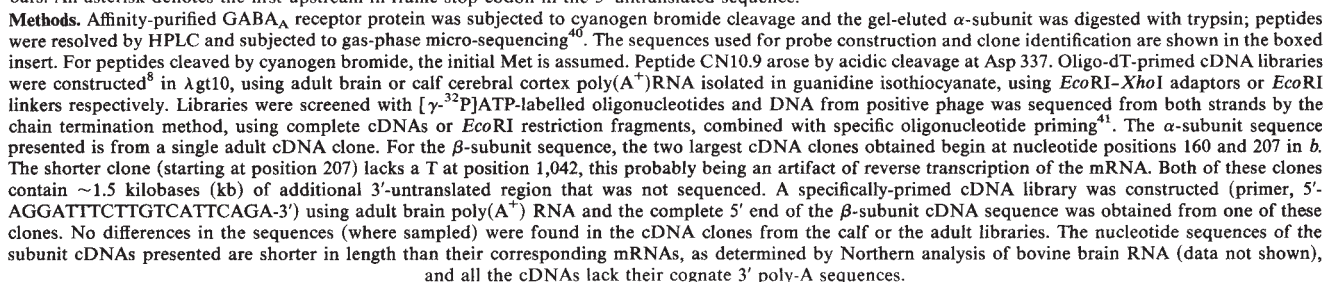
These cDNAs code for polypeptides with deduced protein sizes of 456 amino acids for the α -chain and 474 amino acids for the β -chain. In both polypeptide sequences the N-terminal 25–30 residues constitute typical signal peptides⁹. As no N-terminal sequence was obtained from the purified GABA_A receptor, our assignments for the actual lengths of these signal peptides (Fig. 1) remain tentative. The proposed mature subunit sizes (α -subunit, 429 residues, M_r 48,800; β -subunit 449 residues, M_r 51,400) agree (within the error limits of SDS gel electrophoresis of hydrophobic polypeptides) with the values found⁶ for the deglycosylated GABA_A receptor subunits (α -subunit, 44K; β -subunit, 55K).

Structural interpretations

The deduced protein sequences (for alignment see Fig. 2a) and hydropathy profiles¹⁰ (Fig. 2b) reveal striking similarities in the structure and general architecture of both GABA_A receptor subunits, suggesting their common evolutionary origin. The overall sequence identity of the mature subunits is 35%, while the homology based on conservative substitutions¹¹ rises to 57%. Four hydrophobic sequences of membrane-spanning length (designated as M1–M4) are present in each subunit. A large N-terminal hydrophilic domain, likely to be extracellularly located, contains several potential sites for N-linked glycosylation (2 in the α -subunit, 3 in the β -subunit). Carbohydrate attachment to the receptor is known to occur *in vivo*^{4,6,12}, and would account for the differences between the observed^{4,6} and

‡ To whom correspondence should be addressed.

§ Present addresses: Laboratory of Molecular Neuroendocrinology, ZMBH, Universität Heidelberg, Im Neuenheimer Feld 282, D-6900 Heidelberg, FRG (P.R.S. and P.H.S.); Department of Pharmacology, University of Maryland School of Medicine, 655 West Baltimore Street, Baltimore, Maryland 21201, USA (D.R.B.).



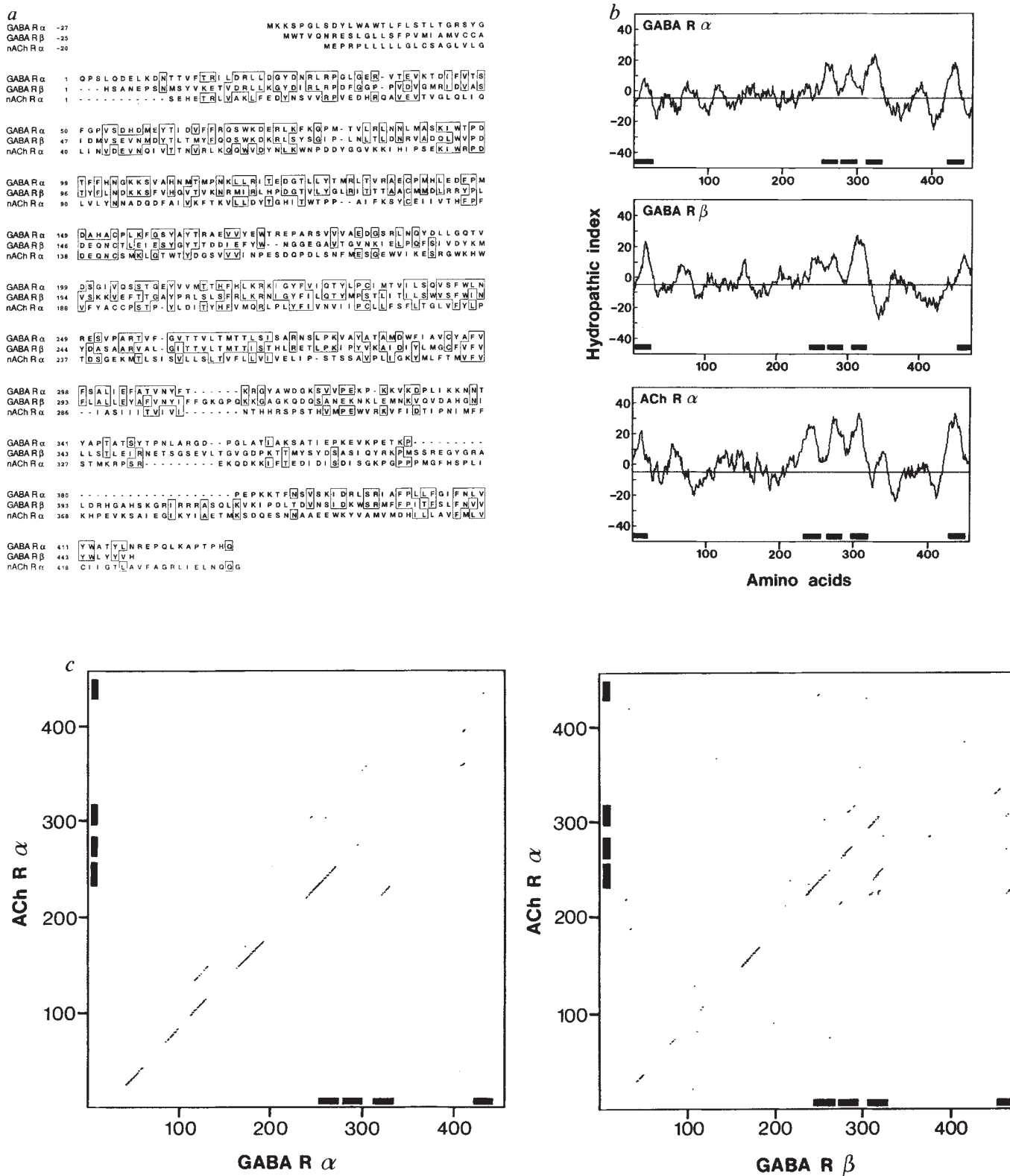


Fig. 2 Homologies of bovine GABA_A receptor α - and β -subunits and the bovine nicotinic acetylcholine receptor α -subunit. **a**, Primary sequence homology: identical residues are boxed. The overall sequence identity between each GABA_A receptor subunit and the bovine nAChR α -subunit is definite but low: 18.5% (α) and 15% (β). **b**, Hydropathy profiles computed according to Kyte and Doolittle¹⁰; window size is 17 residues, plotted with a 1-residue interval. Similar profiles were obtained when the hydropathy index of Hopp and Woods⁴² was used. **c**, Diagonal matrix (DIAGON) conservative homology comparison, computed⁴³ using the mutation data matrix MDM₇₈, of the bovine muscle nAChR α -subunit¹⁴ versus (left) the α -subunit or (right) the β -subunit of the GABA_A receptor. The points correspond to mid-points of 21-residue spans giving a double-matching probability of <0.0005. Sequence regions of strong similarity between the pair by this criterion (including strictly conservative replacements as well as identities) are revealed by unbroken diagonal segments. Very similar plots were obtained when other subunits of the bovine muscle nAChR were used instead of its α -subunit. Solid bars indicate proposed signal sequences (**b**) and transmembrane domains (**b** and **c**).

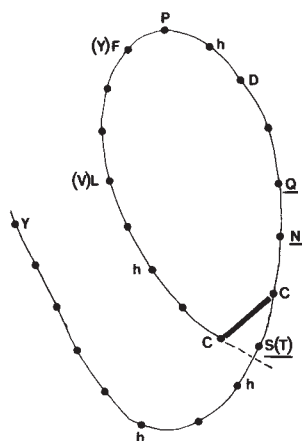


Fig. 3 A loop-containing structure common to the GABA_A and vertebrate nicotinic ACh receptor subunits. This loop is formed by presumed disulphide bonding between Cys 139 and Cys 153 (GABA_A α -subunit numbering). Positions identical (or with one highly conservative variant, as shown) in all nAChR subunits and the β -subunit of the GABA_A receptor are marked with uppercase symbols; h denotes that a hydrophobic residue only is found at that position. A hypothetical interaction between the C-terminal constant Tyr and hydrophobic side-chains of the loop may occur. In the one invertebrate nAChR subunit sequence known²⁰, all of the positions marked are occupied similarly except for 152 (T), 154 (I) and 162 (F) (GABA_A receptor α -subunit numbering). In the bovine GABA_A receptor α -subunit all of this also holds, except that the three residues underlined differ from those in the β -subunit. The disulphide-bridged loop can readily form (in a model) a β -structure, with a hairpin turn centred at the invariant Pro.

predicted molecular weights of the natural and cDNA-deduced GABA_A receptor subunits respectively.

Interestingly, no sequence homology is apparent in the region, presumably intracellularly located, which connects M3 and M4. This region is of different length in the two subunits (87 residues in the α -subunit, 124 in the β -subunit) and contains in the β -subunit a unique cAMP-dependent serine phosphorylation consensus¹³ sequence (Arg-Arg-Arg-Ala-Ser).

Receptor super-family

The domain distribution in the GABA_A receptor subunits is remarkably similar (Fig. 2b) to that seen^{14,15} in the various subunits of the nicotinic acetylcholine receptor (nAChR). All the polypeptides of these two classes are similar in size and contain an identical number and distribution of predicted hydrophobic transmembrane regions. There is some homology between the sequences of either the α - or β -subunits of the GABA_A receptor and any one of the bovine nAChR subunits; this homology is localized to several specific regions of the sequence (Fig. 2c). For the α -subunit of the GABA_A receptor, this is most notable for the M1 region and for some segments in the N-terminal hydrophilic domain, and in these limited regions sequence identity is 34% and homology (conservative¹¹ replacements) is 62%. For the β -subunit the assumed membrane regions M1–M3 show the same levels of homology (Fig. 2c).

As a consequence of these homologies, several structural features emerge as common to both receptor types and may well prove of general functional significance for all ion channel receptors. Most notably, such shared features include specific conformational motifs within the extracellular ligand-binding domain, as well as structural specifications for the gated channel. Thus, a major feature is an extracellular β -structural loop which can be formed by the disulphide bonding of two conserved cysteine residues (positions 139 and 153, GABA_A receptor α -subunit numbering). This loop (Fig. 3), already known to be

formed *in vivo* in the nAChR^{16,17}, plus an adjacent region is now seen to contain three other invariant amino-acids (Pro 147, Asp 149 and Tyr 162) as well as strongly conserved residues at numerous other positions; these features are preserved in all the polypeptides of the GABA_A receptor and the 20 known^{14,15,18} (including neuronal) nAChR subunits. This loop also contains an N-linked glycosylation site which is conserved between the GABA_A β -subunit and the nAChR subunits, and has been shown to be used *in vivo* in the latter^{16,19}.

Although containing channels of mutually exclusive ion selectivity, GABA_A receptor and nAChR subunits share several features of their transmembrane sequences which are presumed to be important for channel function. Without exception, the M1 region in all these subunits contains a proline residue at an identical position and embedded in a conserved protein environment, the consensus sequence being Pro-Cys-X-Leu (where Pro 1 is invariant, Cys 2 is replaced by Ser in the GABA β -subunit or Thr in an invertebrate nAChR subunit²⁰, Leu 4 is replaced by Met in the GABA α -subunit and X is any amino acid). The invariant proline residue is expected to introduce a bend in the transmembrane α -helix. The ensuing protrusion into the channel lumen might keep the channel closed in the absence of neurotransmitter. A second general feature of these channels, regardless of charge preference, is the abundance of threonine and serine residues in the M2 region of all of the subunits: there are 8 of these in M2 in each GABA_A receptor subunit, or 32 per receptor (assuming the $\alpha_2\beta_2$ stoichiometry^{5,6}). These residues may be important in the formation of a hydrophilic lining of the channel, conducive to ion flow. Indeed, there is chemical^{21,22} and mutagenesis²³ evidence on the nAChR which has implicated the M2 segment in the ion channel structure.

Notably lacking in either subunit of the GABA_A receptor are certain features which have been considered of functional importance in the nAChR, in particular a proposed amphipathic helix (MA) located in the latter between the M3 and M4 domains (which has been proposed to line the channel lumen²⁴), and two adjacent cysteines in the extracellular domain at positions 192–193 in every nAChR α -subunit (proposed to be in a channel-gating switch¹⁷). Their absence in the GABA_A receptor indicates that these proposed structures are not a general feature of chemically-gated ion channels.

The structural similarities discussed above, however, strongly indicate that a super-family of chemically-gated ion channel receptors exists. This family is expected to include other amino-acid activated channels such as those gated by glycine or by excitatory amino-acid neurotransmitters. In fact, a concurrent report²⁵ on the sequence of a 48K subunit of the rat spinal cord glycine receptor shows marked similarities to the transmembrane distribution and some of the other structural features discussed above. The presence of both anionic and cationic channel types within this family (including, even, nicotinic receptor anionic channels, as seen in the invertebrate *Aplysia*²⁶) suggests that the integral ion channel and the site for its activation can be discrete and separable domains of these proteins. This super-family can be compared in its generality with that of the G-protein-coupled receptors, which includes the various β -adrenergic, muscarinic and opsin types^{27–29}. In line with their entirely different transduction mechanisms, these two protein super-families are not related at all in either sequence or domain pattern.

Expression in the oocyte system

mRNAs encoding membrane receptors can be translated conveniently and efficiently in the *Xenopus* oocyte, which can fully assemble them in its membrane^{30,31}. This system can produce functional GABA_A receptors when total mRNA from chick or rat brain is micro-injected^{32,33}. To confirm the authenticity of the cloned α - and β -subunit cDNAs as encoding the complete GABA_A receptor, we prepared α -subunit and β -subunit RNAs by inserting the cloned cDNAs into plasmids containing the bacteriophage SP6 promoter and performing *in vitro* transcrip-

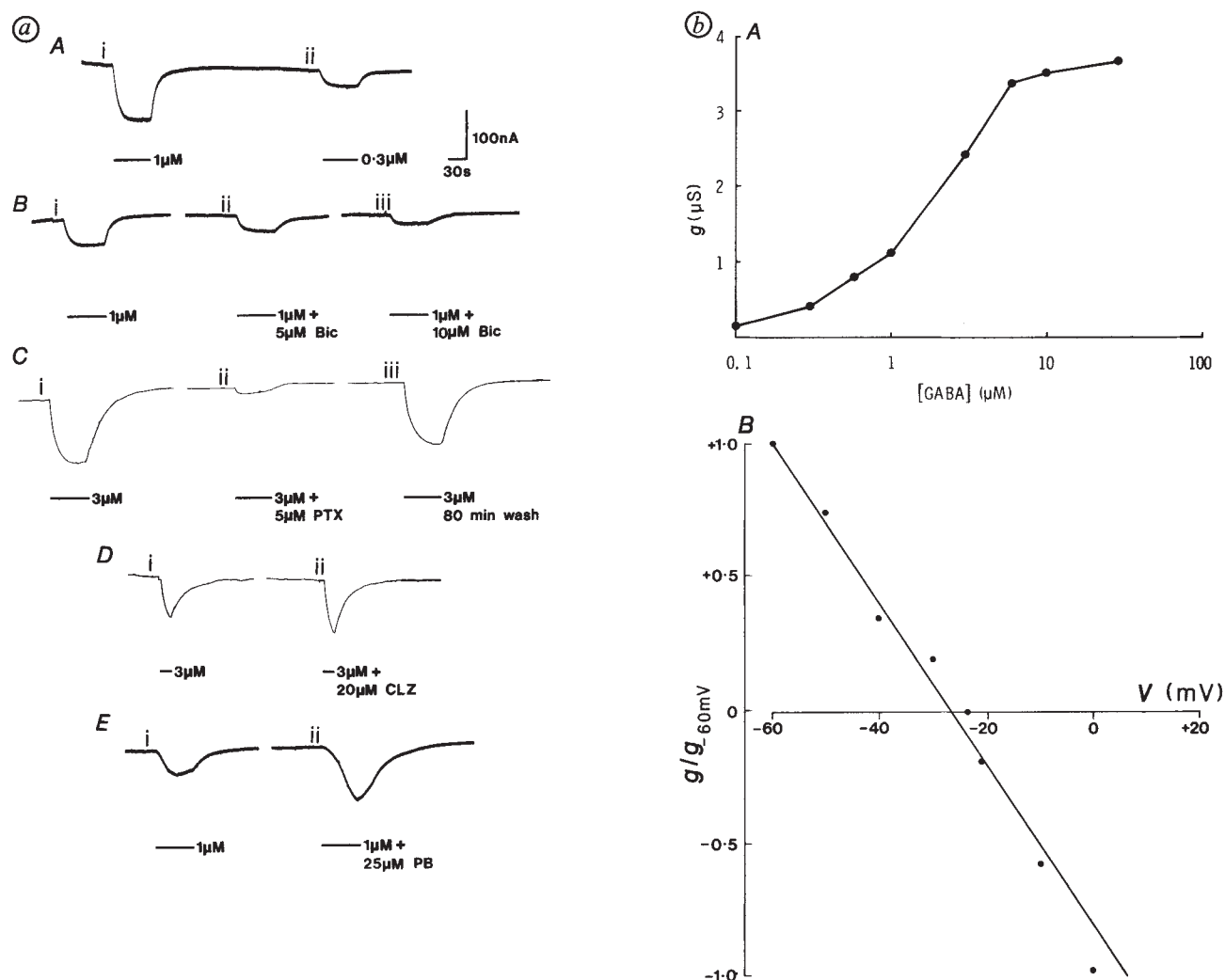


Fig. 4 Expression of the GABA_A receptor in the *Xenopus* oocyte. *a*, GABA responses recorded from individual oocytes co-injected with the α - and β -subunit-specific RNAs derived from the cloned cDNAs. Downward deflections indicate inward currents (see scale); the duration of the GABA application is indicated by the horizontal bar. *A*, Membrane conductance change evoked by 1 μM (i) or 0.3 μM (ii) GABA. *B*, Membrane conductance change evoked by: (i) 1 μM GABA; (ii) 1 μM GABA plus 5 μM bicuculline methobromide (Bic); (iii) 1 μM GABA plus 10 μM bicuculline methobromide. *C*, Membrane conductance change evoked, in another RNA-injected oocyte, by: (i) 3 μM GABA; (ii) 3 μM GABA plus 5 μM picrotoxin (PTX); (iii) 3 μM GABA, after washing out the picrotoxin for 80 min. *D*, Membrane conductance changes in response to: (i) 3 μM GABA; (ii) 3 μM GABA plus 20 μM chlorazepate (CLZ). The latter was present also for 10 min before GABA application. *E*, Membrane conductance changes evoked by: (i) 1 μM GABA; (ii) 1 μM GABA plus 25 μM (–) pentobarbital (PB). Note that no response to GABA was ever observed in numerous experiments (not illustrated) in which either the α - or β -subunit-specific RNA was injected alone. *b*, Quantitative analyses of oocyte expression. *A*, GABA log-dose/conductance relationship. Ordinate, change in peak membrane conductance (g , in μS); abscissa, GABA concentration (μM , log scale). Typical data obtained from a single oocyte co-injected 48 h previously with α - and β -subunit-specific RNAs. *B*, Measurement of the reversal potential for GABA (used at 1 μM) in a single oocyte, co-injected with α - and β -subunit-specific RNAs. The ordinate is g , expressed as a fraction of g measured at -60 mV. The reversal potential is -27.5 mV. **Methods.** The entire α -subunit cDNA was subcloned by its *EcoRI* adaptor sequence into pSP65 (ref. 34) (pbGR α sense). The β -subunit cDNA was constructed by using a *HindIII* (restriction site in M13 polylinker)–*EagI* fragment from the specifically-primed cDNA clone and an *EagI*–*BglII* fragment from the longer of the two oligo(dT)-primed cDNA clones. Both the *HindIII* and *BglII* sites were filled in with T4 DNA polymerase and the fragments ligated at the *EagI* site. The ligation product was gel-eluted and ligated to pSP65 (pbGR β sense). RNA transcripts were obtained after cleavage of both template DNAs with *HindIII* and transcribing *in vitro* with SP6 RNA polymerase³⁴. Transcripts were capped with m⁷G(5')ppp(5')G. The synthetic RNA was injected into *Xenopus* oocytes (each subunit-specific RNA concentration was 0.25 ng nl⁻¹; total volume injected per oocyte, ~30 nl); the oocytes were then incubated at 19 °C for 2 to 4 days in modified Barth's medium⁴⁴. All recordings were made at 20–23 °C in amphibian Ringer solution⁴⁵. GABA-evoked currents were recorded by a two micro-electrode voltage-clamp technique at -60 mV holding potential, with superfusion where indicated of GABA and other drugs⁴⁵. Chord conductances were calculated, where g is shown.

tion³⁴. The two pure RNAs produced were micro-injected, singly or together, into *Xenopus* oocytes and pulses of GABA were subsequently superfused over the oocyte while recording membrane currents under voltage-clamp conditions. Only when both RNAs were injected together was any current response to GABA application observed (Fig. 4a), indicating correct functional assembly. This response was large and immediate at, for example, 0.3 or 1 μM GABA concentration, and (at a holding

potential of -60 mV), the threshold for an observable current change was in the range of 10^{-7} M. The peak inward current was reproducibly dose-dependent (Fig. 4b); the half-maximal GABA concentration was 1.8 μM , noticeably lower than that (30–60 μM) observed with rat³³ or chick³² brain mRNA. Elevated expression from the pure mRNAs may account for the reduced desensitization, relative to that seen after rat brain mRNA injection³³, observed with a 1-min GABA application.

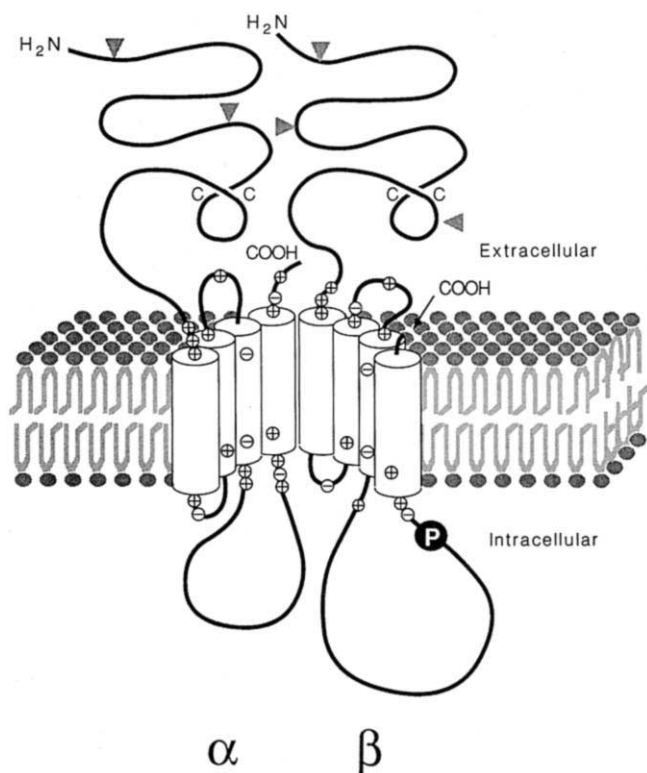


Fig. 5 A schematic model for the topology of the GABA_A receptor in the cell membrane. Four membrane-spanning helices in each subunit are shown as cylinders. The structure in the extracellular domain is drawn in an arbitrary manner, but the presumed β -loop (Fig. 3) formed by the disulphide bond predicted at cysteines 139 and 153 is shown. Potential extracellular sites for *N*-glycosylation are indicated by triangles. Those charged residues which are located within or close to the ends of the membrane domains are indicated as small circles with charge marked. The site for cAMP-dependent serine phosphorylation, present only in the β -subunit, is marked by an encircled P. It is proposed that two copies of each of these subunit structures are complexed in the receptor molecule so as to align the membrane-spanning domains, only some of which will form the inner wall of a central ion channel.

This desensitization may plausibly be modulated by serine phosphorylation at the cAMP-dependent site found in the β -subunit; such a system could become overloaded at an exceptional receptor density. GABA responsiveness was totally absent in appropriate controls or in oocytes injected with the α - or the β -subunit RNA alone.

The current-voltage curve for the RNA-induced GABA response was regular in form and showed a reversal potential of -27.5 mV (Fig. 4b), which corresponds to the equilibrium potential of chloride in the *Xenopus* oocyte³⁵. This evidence, plus recent further work by E. Levitan and E. A. Barnard (to be published) showing that the extracellular chloride concentration dependence and other properties are as predicted, demonstrates that an anion channel, formed by the injected RNAs, is opened by GABA in the oocyte membrane, as at brain synapses. Although the non-injected oocyte can also produce, upon activation of its intrinsic muscarinic receptors, a fast inward chloride current³⁵, the channel induced here is a new and different one, as the two currents, quite apart from their entirely different controlling pharmacologies, are unlike in their voltage dependencies and other channel characteristics (to be described elsewhere).

The pharmacology of the oocyte-expressed receptor was compared with that of the native GABA_A receptor^{1,2,36}. Current flow

in the oocyte-expressed receptor was strongly blocked (Fig. 4a) by either bicuculline, a highly-selective competitive antagonist¹ or by picrotoxin, a convulsant blocking agent of the GABA-activated channel^{1,36} which shows slow binding reversibility. As expected, the blockade by bicuculline was readily reversible upon wash-out (not shown), whereas that of picrotoxin was only slowly reversible (Fig. 4a). A benzodiazepine (chlorazepate, at 20 μ M) and barbiturates (for example, (-)pentobarbital at 25 μ M) allosterically potentiated the GABA response, the latter more strongly. Benzodiazepines are known to achieve their potentiation by increasing the frequency of channel opening, whereas barbiturates do so by prolonging the channel lifetime^{1,36}.

Quantitative comparison of the responses observed with GABA and its modulators cannot be made in detail to those on native (neuronal) receptors, as the electrophysiology of bovine neuronal GABA receptors has not been reported. However, the evidence presented here establishes that the α - and β -subunits are necessary and sufficient to form the functional receptor for GABA with its integral chloride channel. The characteristic binding sites of the GABA_A receptor complex, as known in brain membranes or in the isolated protein, are all formed by the transcription of the two cloned cDNAs and the assembly of the subunits. Moreover, allosteric potentiation of the channel opening by barbiturates and by benzodiazepines is present, as is true for the native receptor.

Structural model

In the light of the similarities of the GABA_A receptor and the nAChR, we propose the following model for the structure and function of the GABA_A receptor complex (Fig. 5). In both subunits the four hydrophobic domains transverse the membrane. These 16 helices must contribute to or stabilize the walls of the channel. Some of these must be placed so as to form the 5.6 Å-bore lumen of this channel^{37,38}. The N-termini are assigned to the extracellular side (as is generally found with receptors which contain signal peptides). This topology would place the C-termini just outside the extracellular surface and locates the cAMP-dependent phosphorylation site in the interior, where it must be if functional. We cannot exclude at this stage, however, a variant of this model, in which the M4 helix lies in the membrane surface (as has been established crystallographically for the reaction centre protein in the chloroplast membrane³⁹), allowing the C-terminus to be intracellular.

The lining of the channel is postulated to be rendered conducive to ion flow by the presence of threonine and serine side-chains, largely from the M2 helix (as discussed above), as well as by the two basic and two acidic residues (probably paired) within the assigned transmembrane domains.

A marked clustering of positively-charged side-chains occurs at both ends of the postulated membrane-spanning domains (Fig. 5). These charged residues are presumed to form the channel mouth. In an $\alpha_2\beta_2$ receptor complex, there would be 26 positive charges and only 4 negative charges on the extracellular side. Most of these clustered basic residues are absent or become of opposite charge in the corresponding nAChR locations. Such clusters of positive charge in the mouth of the channel would act as an anion-selective filter and increase the driving force for chloride flow upon opening of the channel gate. This model is in agreement with the channel reversibility behaviour³⁷, channel conductance studies which suggest that at least two anion binding sites are associated with the channel³⁷, and the strong positive influence of halide ion concentration on the binding of some of the specific ligands for this receptor¹.

On a more speculative note, the gating mechanism may be operated through a dual configuration of the M1 sequence within both subunits. A flexure in M1 is introduced by the unique proline residue noted to be in an invariant motif; M1 could, therefore, be repositioned upon neurotransmitter binding to open the full channel lumen. The binding sites for the various

GABA and benzodiazepine agonists and antagonists are probably located in the large N-terminal extracellular domains, which contain a total of 10 potential N-glycosylation sites per receptor complex. Such binding site locations are supported by the finding¹² that deglycosylation of the rat GABA_A receptor modifies the binding of benzodiazepine agonists and antagonists. Thus, GABA binding to the β -subunits⁵ of the receptor complex would induce a conformational change, exposing some of the positively-charged residues at the channel mouth and possibly shifting the configuration of M1, resulting in chloride ion flux. Allosteric modulation of these conformational changes must be produced by the binding of benzodiazepines or barbiturates. The region linking M3 and M4, which shows no homology between the α - and β -subunits and contains the phosphorylation

site, is a potential location for other sites, as yet unknown, mediating intracellular control of channel activity.

The testing of such a model, and the establishment of the mechanisms of channel function and receptor modulation, should be feasible now by means of mutational analysis of the cloned α - and β -subunit cDNAs, coupled with single channel recording in the oocyte expression system.

N.F. holds an EMBO Long-term Fellowship, D.R.B. held a Fogarty Senior International Fellowship and F.A.S. holds a Royal Society University Research Fellowship. We thank Michael Squire (Cambridge) for expert technical help with some of the DNA sequencing, Carole Morita (Genentech) for help with computer graphics and Mary Wynn (Cambridge) for word processing.

Received 26 May; accepted 16 June 1987.

- Olsen, R. W. & Venter, J. C. (eds) *Benzodiazepine/GABA Receptors and Chloride Channels: Structural and Functional Properties* (Liss, New York, 1986).
- Turner, A. J. & Whittle, S. R. *Biochem. J.* **209**, 29–41 (1983).
- Sigel, E., Stephenson, F. A., Mamalaki, C. & Barnard, E. A. *J. biol. Chem.* **258**, 6965–6971 (1983).
- Sigel, E. & Barnard, E. A. *J. biol. Chem.* **259**, 7219–7223 (1984).
- Casalotti, S. O., Stephenson, F. A. & Barnard, E. A. *J. biol. Chem.* **261**, 15013–15016 (1986).
- Mamalaki, C., Stephenson, F. A. & Barnard, E. A. *EMBO J.* **6**, 561–565 (1987).
- Schoch, P. *et al. Nature* **314**, 168–171 (1985).
- Huynh, T. V., Young, R. A. & Davis, R. W. in: *DNA Cloning: A Practical Approach*, Vol. 1 (ed. Glover, D. M.) 49–78 (IRL, Oxford, 1985).
- Von Heijne, G., *Nucleic Acids Res.* **14**, 4683–4690 (1986).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Schwartz, R. M. & Dayhoff, M. O. in *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3 (ed. Dayhoff, M. O.) 353–358 (Nat. Biomed. Res. Fdn., Washington DC, 1978).
- Sweetnam, P. M. & Tallman, J. F. *Molec. Pharmacol.* **29**, 299–306 (1986).
- Feramisco, J. R., Glass, D. B. & Krebs, E. G. *J. biol. Chem.* **255**, 4240–4245 (1980).
- Kubo, T. *et al. Eur. J. Biochem.* **149**, 5–13 (1985).
- Popot, J.-L. & Changeux, J.-P. *Physiol. Rev.* **64**, 1162–1239 (1984).
- Criado, M., Sarin, V., Fox, J. L. & Lindstrom, J. *Biochemistry* **25**, 2839–2846 (1986).
- Kao, P. N. & Karlin, A. J. *J. biol. Chem.* **261**, 8085–8088 (1986).
- Goldman, D. *et al. Cell* **48**, 965–973 (1987).
- Conti-Tronconi, B. M., Hunkapiller, M. W. & Raftery, M. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2631–2634 (1984).
- Hermans-Borgmeyer, I. *et al. EMBO J.* **5**, 1503–1508 (1986).
- Hucho, F., Oberthür, W. & Lottspeich, F. *FEBS Lett.* **205**, 137–142 (1986).

- Giraudat, J., Dennis, M., Heidmann, T., Chang, J.-Y. & Changeux, J.-P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2719–2723 (1986).
- Imoto, K. *et al. Nature* **324**, 670–674 (1986).
- McCarthy, M. P. *et al. A. Rev. Neurosci.* **9**, 383–413 (1986).
- Grenningloh, G. *et al. Nature* **328**, 215–220 (1987).
- Ono, J. K. & Salvaterra, P. M. *J. Neurosci.* **1**, 259–270 (1981).
- Dixon, R. A. F. *et al. Nature* **321**, 75–79 (1986).
- Kubo, T. *et al. Nature* **323**, 411–416 (1986).
- Nathans, J. & Hogness, D. S. *Cell* **34**, 807–814 (1983).
- Sumikawa, K., Houghton, M., Emtage, J. S., Richards, B. M. & Barnard, E. A. *Nature* **292**, 862–864 (1981).
- Barnard, E. A., Miledi, R. & Sumikawa, K. *Proc. R. Soc. Lond. B* **215**, 241–246 (1982).
- Smart, T. G., Constanti, A., Bilbe, G., Brown, D. A. & Barnard, E. A. *Neurosci. Lett.* **40**, 55–59 (1983).
- Houamed, K. M. *et al. Nature* **310**, 318–321 (1984).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
- Dascal, N., Landau, E. M. & Lass, Y. *J. Physiol.* **352**, 551–574 (1984).
- Study, R. E. & Barker, J. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7180–7184 (1981).
- Bormann, J., Hammill, O. P. & Sakmann, B. *J. Physiol.* **385**, 243–286 (1987).
- Unwin, P. N. T. *Nature* **323**, 12–13 (1986).
- Deisenhofer, J., Epp, O., Miki, P., Huber, R. & Michel, H. *Nature* **318**, 618–624 (1985).
- Rodriguez, H. *J. Chromatog.* **350**, 217–225 (1985).
- Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309–321 (1981).
- Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824–3828 (1981).
- Staden, R. *Nucleic Acids Res.* **10**, 2951–2961 (1982).
- Barnard, E. A. & Bilbe, G. in *Neurochemistry: A Practical Approach* (eds Turner, A. J. & Bachelard, H.) 243–270 (IRL, Oxford, 1987).
- Van Renterghem, C. *et al. Molec. Brain Res.* **2**, 21–31 (1987).

LETTERS TO NATURE

Warming of Miranda during chaotic rotation

Robert Marcialis & Richard Greenberg

Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

Miranda, a satellite of Uranus, appears in Voyager images to have had an active geological history, seemingly characterized by relaxation of very large-scale topography or the sinking of large blocks of material accreted by the satellite, with associated extrusion, eruption, and flow of the icy material on the body¹. The low temperature of Miranda (<100 K) would preclude such mobilization of water ice. Hypothetical methane or ammonia constituents have been invoked to circumvent the problem². Here we offer an alternative explanation: topography or anomalous blocks themselves may once have provided the conditions for heating and softening the supporting material, even pure ice, by permitting chaotic rotation of the satellite (J. Wisdom, unpublished lecture). Moreover, this heating mechanism is self-limiting due to rapid damping of the orbital eccentricity. Topography at the presently observed scale (roughly 20 km on this 240-km-radius body) could be preserved.

Chaotic behaviour can occur when resonances are so strong that they overlap³. A satellite that rotates synchronously (like the Moon) is in a 1/1 resonance with its orbital motion. This behaviour is stable if the satellite is elongated or has a density asymmetry. Such rotation is maintained by the torque of the planet on the asymmetrical mass distribution. During oscillations about exact synchronicity, the rotation rate can vary to an

extent limited by the amount of elongation of the satellite. If the elongation is sufficient, the rate may approach other important resonance rates, such as 3/2 the orbital angular velocity. This case would be an example of an overlapping resonance. Such a non-spherically symmetric, prolate satellite would probably experience chaotic, unpredictable changes in its rotation rate given a wide range of initial conditions.

The criterion for such resonance overlap can be expressed in terms of the principal moments of inertia, $A < B < C$. The quantity $(B - A)/C$ must be greater than ~0.1 (ref. 4), which it is for even a slightly prolate body or slightly asymmetric density distribution. For example, an elongation of only about 5% would be sufficient assuming uniform density distribution. Thus, any satellite with moderate asphericity, perhaps due to a high-velocity impact or low-velocity accretion of large components, would be a candidate for chaotic rotation.

Shape is not the only criterion which must be satisfied for chaotic rotation. Additionally, the satellite's orbit must be significantly eccentric in order for spin-orbit resonances other than the 1/1 to be important. If the eccentricity e is too small, a resonance like 3/2 would be too weak to interfere, even if the 1/1 overlaps it. Criteria for adequate e values have not yet been quantified. Hyperion, with $e = 0.1$, almost certainly rotates chaotically³.

Miranda's geology suggests that the satellite was once non-spherical, whether due to accretion of large blocky constituents⁴ or a major impact event. If its orbit were eccentric, it may have rotated chaotically. A non-synchronously rotating satellite has its body worked by tides and is thus heated at the rate

$$dE/dt = 3GM^2(n/a)(k_2/Q)(r/a)^5 \quad (1)$$

where M is the mass of the planet and n , a , k_2 , Q and r are

the satellite's mean motion, orbital semi-major axis, Love number, dissipation parameter, and radius, respectively⁵. This formula applies when the rotation rate is of the same order of magnitude as n . For synchronous rotation, tidal heating would be much slower, by a factor of e^2 . Chaotic rotation is non-synchronous, so equation (1) applies (J. Wisdom, unpublished lecture).

To evaluate this heating rate, we need an estimate of k_2 . For a small rigid satellite like Miranda,

$$k_2 = 3\rho g r / (19\mu) \quad (2)$$

where ρ is the density 1.2 g cm^{-3} , g is the surface gravity 9 cm s^{-2} and μ is the rigidity. Following the estimates of Squyres *et al.*⁶ for Enceladus, we adopt values $\mu \sim 4 \times 10^{10} \text{ dyn cm}^{-2}$ and $Q \sim 20$, to obtain $k_2 \sim 10^{-3}$ and $dE/dt \sim 6 \times 10^{18} \text{ erg s}^{-1}$. This rate is 10^4 times more than expected due to radiogenic heating alone⁷ and ten times more intense per gram than the present tidal heating of Io. Continued heating at such a rate would probably lead to internal softening and relaxation toward a hydrostatic figure, possibly accompanied by surface flows.

This intense heating could not continue for long. The energy must ultimately come out of the satellite's orbit. To conserve angular momentum, the eccentricity must decrease at the rate

$$de/dt = dE/dt / (mn^2 a^2 e) \quad (3)$$

where m is the satellite's mass. Thus e must plunge to zero in time $t \sim 10^8 e_0^2 \text{ yr}$, where e_0 is the original value of e . Adopting a reasonable value of $e_0 = 0.1$, we find that $t \sim 10^6 \text{ yr}$. Because non-zero orbital eccentricity is required to maintain chaotic rotation, the episode of heating terminates in $\sim 1 \text{ Myr}$.

The total heat dissipated in Miranda during that period would have been $2 \times 10^{32} \text{ erg}$. With a heat capacity C_p of $\sim 1.3 \times 10^7 \text{ erg g}^{-1} \text{ K}^{-1}$, the temperature would rise by $\sim 200 \text{ K}$ above the surface temperature of 70 K , assuming uniform heating in the body. The temperature rise could not be so great near the surface where conduction removes heat as it is deposited. At the end of the million-year heating period, this cooler boundary layer would have a thickness

$$d \sim (\kappa t)^{1/2} \sim 25 \text{ km} \quad (4)$$

with thermal diffusivity $\kappa \sim 0.15 \text{ cm}^2 \text{ s}^{-1}$ for ice. The temperature profile would range from 70 K at the surface to the 200 K higher value at 25 km and deeper. Solid-state convection would not occur in this transition layer because at these temperatures the effective viscosity η of ice ($\sim 10^{20} \text{ g cm}^{-1} \text{ s}^{-1}$; see refs 2, 8) is too great and the layer is too thin⁹.

The relatively cool upper 25 km has a relaxation time of about 10^6 yr for topography with a horizontal length scale (λ) of 20 km using the above parameters in the following¹⁰:

$$t = 2\pi\eta / (\rho g \lambda) \quad (5)$$

However, the timescale for substantial cooling in this layer is also about 10^6 yr so the rate of relaxation of 20 km (or smaller) topography would have decreased before much degradation could occur.

For larger structure (greater λ), such as the global non-hydrostatic figure that permitted chaotic rotation, the relaxation timescale is shortened by two factors: first, the inverse λ dependence in equation (5) and second, an additional factor that accounts for the much-reduced viscosity^{2,8} at the higher temperature below 25 km . This second factor is unity for $\lambda < d$, but is 0.5 for $\lambda = 100 \text{ km}$ and 0.1 for λ near 240 km , the radius of the satellite¹⁰. Thus global scale topography would have relaxed away before the satellite cooled down from its chaotic rotation heating. Numerical integration of the heat flow equation shows that all vestiges of the implanted heat will be dissipated within 100 Myr .

This explanation of the heat source for geological activity, the relaxation of global structure, and retention of significant

local topography does not require any hypothetical constituents added to the water ice. In contrast, Stevenson and Lunine² hypothesize an intergranular cryogenic fluid such as methane to obtain lower viscosities so that deformation can occur at low temperatures. Similarly, Squyres *et al.*⁶ had to assume a eutectic mix of water and ammonia in order to bring about even marginal melting of Enceladus.

Our calculation of the total heat dissipated during chaotic rotation, and of the corresponding internal temperature rise is independent of the 30% uncertainty in Miranda's mass¹ and of the much greater uncertainty in the value of k_2/Q . We have adopted the same minimal value for Q as Squyres *et al.*⁶. A larger value of Q would slow the heating rate, but lengthen the chaotic era, preserving the total energy deposited. The depth of the conduction layer increases, but only as $Q^{1/2}$. Also, any effect of a larger Q would be counteracted by an increase in the value of k_2 that would have accompanied the warming and softening of the interior.

The total energy dissipated is very sensitive to the initial eccentricity e_0 , because the duration of the chaotic rotation is proportional to e_0^2 . For our calculations, we have assumed a plausible value $e_0 \sim 0.1$ and demonstrated that the appropriate amount of heating could take place. Such a value could have been primordial. It is comparable to the present orbital inclination in radians¹. Perhaps e_0 was generated by some past orbital resonance, such as a resonance with Ariel that may have been traversed during tidal evolution of the satellites' orbits¹¹. If e_0 were a few times smaller, the heating of Miranda would have been greatly reduced. In that case, the effects we describe might still have occurred given a somewhat smaller heat capacity, or the presence of impurities such as methane or ammonia which would permit equivalent softening with less heat.

If the heating process described here were active before the cratering and basin-forming bombardment of Miranda's surface, all geological evidence would have been destroyed. Certainly any geological expression of accretional heating (which by definition occurred early and which was at most only marginally adequate to have significant effect) must have been erased. On the other hand, it is plausible that chaotic rotation and heating took place later on. The last basin-forming events appear to have occurred after formation of the cratered terrain¹². If these large impacts gave the modest asymmetry needed for chaotic rotation, our mechanism could have heated Miranda after most craters were in place. These ideas will be tested as a complete geological chronology emerges.

It is possible that some or all of the other icy satellites of comparable size to Miranda experienced similar chaotic rotation if their figures were originally comparably irregular³. Repeating the above calculation for the inner satellites of Saturn and for the other satellites of Uranus shows that they could have been heated comparably to Miranda. With the same assumed initial eccentricity, we find they generally would have received several times more heat per volume than Miranda. They would have been overheated from the point of view of retaining even local topographic structure. Irregularities in global shape surely would have been eliminated and the interiors differentiated. Melt flows on the surface might have been created, and then possibly erased, depending on the impact chronology in each case. For those satellites whose eccentricities were always small, tidal heating would have been negligible even during chaotic rotation. Today, only Hyperion remains in chaotic rotation. Its resonance with Titan maintains the eccentricity, but at its great distance from Saturn it does not heat up very much. According to our model, Miranda's eccentricity damped away allowing just enough heating for global relaxation and geological activity, yet allowing local topography to be retained.

We thank J. A. Burns, S. Croft, J. Lunine, H. J. Melosh, S. J. Peale and J. Wisdom for helpful discussions and review of this manuscript. Croft provided the estimates of thermal diffusivity and heat capacity used in our calculations.

Received 12 January; accepted 24 April 1987.

1. Stone, E. C. & Miner, E. D. *Science* **233**, 39–43 (1986).
2. Stevenson, D. J. & Lunine, J. I. *Nature* **323**, 46–48 (1986).
3. Wisdom, J., Peale, S. J. & Mignard, F. *Icarus* **58**, 137–152 (1984).
4. Kerr, R. A. *Science* **235**, 31 (1987).
5. Hubbard, W. B. *Planetary Interiors* 96–102 (Van Nostrand Reinhold, New York, 1984).
6. Squyres, S. W., Reynolds, R. T., Cassen, P. M. & Peale, S. J. *Icarus* **53**, 319–331 (1983).
7. Schubert, G., Spohn, T. & Reynolds, R. T. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 224–292 (University of Arizona Press, Tucson, 1986).
8. Poirer, J. P. *Nature* **299**, 683–688 (1982).
9. Cassen, P. M. & Reynolds, R. T. *Geophys. Res. Lett.* **6**, 121–124 (1979).
10. Melosh, H. J. *Impact Cratering: A Geological Process* (Oxford University Press, in the press).
11. Peale, S. J. *Bull. Am. astr. Soc.* **18**, 785 (1986).
12. Croft, S. K. *Lunar planet. Sci.* **XVIII** 207–208 (1987).

Experiments on laser guide stars at Mauna Kea Observatory for adaptive imaging in astronomy

Laird A. Thompson* & Chester S. Gardner†

* Institute of Astronomy, University of Hawaii at Manoa, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1406 West Green Street, Urbana, Illinois 61801, USA

Atmospheric turbulence severely limits the resolution of ground-based astronomical telescopes. Under good seeing conditions at the best observatory sites, resolution at visible wavelengths is typically limited to about 1 arc s. During the past 15 years, adaptive optical systems with electrically deformable mirrors have been developed to compensate for turbulence^{1,2}. Unfortunately, these systems require bright reference sources adjacent to the object of interest and can only be used to observe the brightest stars. Foy and Labeyrie³ were the first to suggest that lasers could be used to create artificial guide stars that might be suitable in controlling an adaptive imaging system. We have recently extended this concept in two ways. First, we have identified the key engineering parameters that optimize the performance of a laser-guided imaging system. Second, on the nights of 21 and 22 January 1987, we conducted experiments at the Mauna Kea Observatory on the island of Hawaii to test the feasibility of using a laser to generate an artificial guide star in the mesospheric sodium layer. Here we describe both the engineering calculations and the results of our first experiments.

The angular resolution of a ground-based telescope is $1.22\lambda/r_0$, where λ is the optical wavelength and r_0 is the diameter of the turbulence seeing cell. At the best observatory sites, under good seeing conditions, $r_0 = 10$ –20 cm, and at a few outstanding sites, such as Mauna Kea, r_0 can exceed 40 cm on rare occasions. If the diameter of the telescope is less than r_0 , a diffraction-limited image will be focused onto the image plane⁴. However, tilt caused by turbulence makes the image dance about so that long-term exposures are blurred. If the object is bright enough, a dynamic tip-tilt mirror can be used to stabilize the image⁵.

When the telescope diameter (D) is larger than r_0 , each seeing cell within the pupil generates a subimage of angular radius, $1.22\lambda/r_0$, that dances in the image plane independently of the subimages from every other seeing cell. Modern adaptive telescopes use a wavefront sensor consisting of a lenslet array followed by a quadrant detector array to measure the centroid positions of each seeing cell subimage. This information is then used to control the deformable mirror^{1,2}. A quadrant detector can determine a subimage centroid with an angular accuracy given approximately by

$$\frac{0.61\lambda/r_0}{\sqrt{N}} \quad (1)$$

where N is the subimage photon count during the measurement

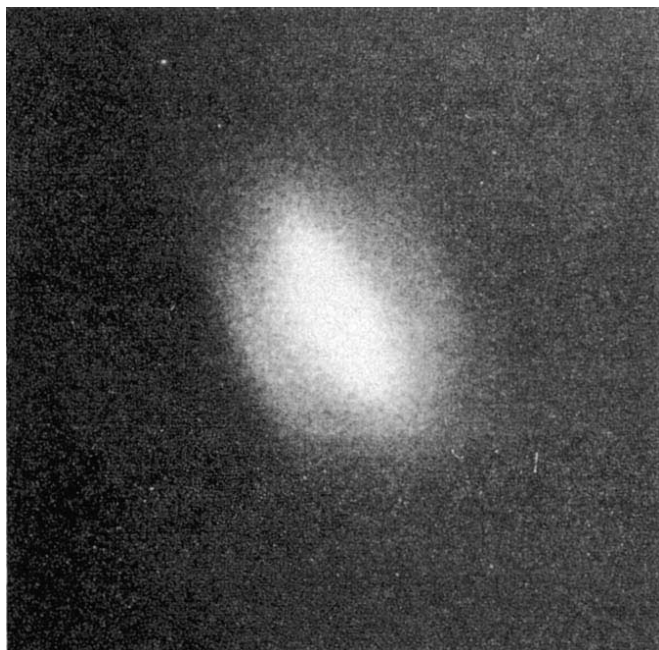


Fig. 1 Image of the laser guide spot obtained on 21 January 1987; exposure time was 8 min starting at 13:43 UT. The image is asymmetric because of the laser beam divergence, imaging geometry and the finite thickness of the sodium layer. The laser beam enters the layer from the top left and leaves towards the lower right. The telescope was fixed near zenith so that background stars drifted through the field of view during the exposure. Two such trails cut horizontally across the image. North is to the top and east to the left.

integration period. The integration time is limited to the turbulence coherence time which is typically ~ 10 ms. To achieve diffraction-limited performance, each subimage centroid must be determined with an accuracy sufficient to reduce the wavefront tilt to $\sim 1/3$ of the residual wavefront distortion remaining in a perfectly tilt-corrected seeing cell (namely, $\sim \lambda/17$).

$$\frac{0.61\lambda/r_0}{\sqrt{N}} \approx \frac{1}{3} \frac{4}{r_0} \frac{\lambda}{17} \approx 0.8 \frac{\lambda}{r_0} \quad (2)$$

Thus, the required photon count at each quadrant detector is $N \approx 60$. To reach the diffraction limit of any ground-based telescope, the required signal levels are very large. Because very few interesting astronomical objects are sufficiently bright or have a suitable guide star located near them, adaptive imaging systems have found only limited use in astronomy.

Laser-guide stars can be generated in the upper atmosphere by either Rayleigh scattering by air molecules in the stratosphere or resonant scattering by alkali metals in the mesosphere. By calculating the laser power requirements for each technique, we find that resonant scattering by sodium atoms is the preferred approach. The chemistry and dynamics of the sodium layer have been studied extensively since the late 1960s with lidar techniques^{6,7}. Meteoric ablation is generally regarded as the dominant source of all mesospheric alkali metals including sodium. The sodium layer is typically confined to the region between 80 and 110 km with a peak near the mesopause at 90 km, where the density ranges from $\sim 10^3$ – 10^4 cm⁻³. The sodium column abundance at mid-latitudes in the Northern Hemisphere varies from a summer minimum of $\sim 2 \times 10^9$ cm⁻² to a winter maximum of $\sim 10^{10}$ cm⁻² in December and January⁶. The seasonal and geographical variations in sodium abundance are now believed to be related to changes in the mesopause temperature that affect the reaction rates of the main chemical loss processes for sodium. For adaptive imaging, sodium abundance is an important parameter because the brightness of a guide

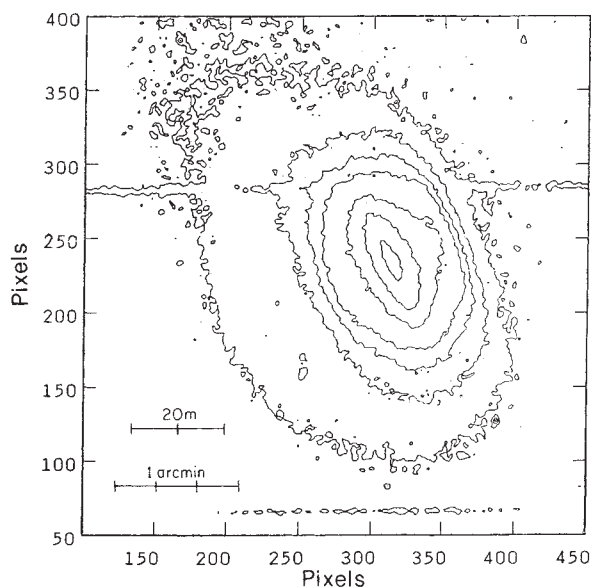


Fig. 2 Photon-count contour plot of the laser-guide spot image. To produce the plot, the image in Fig. 1 was first smoothed by a 3×3 pixel gaussian filter. The contours are 1,300, 1,000, 800, 600, 500, 400 and 310 counts per pixel. The peak signal count was 1,080 and the average background count was 280. The pixel size was $3.4 \mu\text{rad}$. The metre scale refers to the spot dimensions at the layer centroid height of 95.1 km.

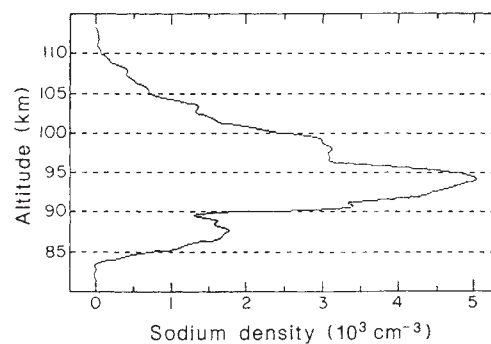


Fig. 3 Sodium density profile measured during the same time period in which the image in Fig. 1 was recorded. The layer column abundance was $5.23 \times 10^9 \text{ cm}^{-2}$, the centroid height was 95.1 km MSL and the r.m.s. thickness was 5.1 km. The structure in the layer is caused by low-frequency internal gravity waves¹¹.

star generated in the layer will be proportional to the column abundance.

To be most effective, the angular radius of the laser-guide star should be equal to $1.22\lambda/r_0$. In this case, the required signal level is given by equation (2). The expected signal level in each subimage can be calculated from the lidar equation⁷

$$N \approx \eta T_A^2 \frac{r_0^2 \lambda E \sigma_{\text{eff}} C_s}{16 z_s^2 h c (1 + \tau_n / \tau_s)} \quad (3)$$

where η is efficiency of telescope and detector (~ 0.075), T_A is one-way transmittance of the lower atmosphere, λ is optical wavelength ($0.589 \times 10^{-6} \text{ m}$), E is laser energy per pulse (J), h is Planck's constant ($6.63 \times 10^{-34} \text{ J s}$), c is velocity of light ($3 \times 10^8 \text{ m s}^{-1}$), z_s is sodium layer centroid height ($\sim 9 \times 10^4 \text{ m}$), σ_{eff} is effective sodium backscatter cross section (m^2), C_s is sodium column abundance (m^{-2}), τ_n is natural lifetime of sodium D_2 excited state ($1.61 \times 10^{-8} \text{ s}$), and

$$\tau_s \approx \frac{0.74 \pi z_s^2 \lambda h c \tau_L}{T_A \sigma_{\text{eff}} r_0^2 E} \quad (4)$$

The laser pulse length is τ_L . When the sodium layer is illuminated by a high-intensity, well-collimated beam, stimulated emission can substantially decrease the number of excited-state atoms available for isotropic fluorescence⁷. The saturation time τ_s is the characteristic time of occurrence of stimulated emission. When τ_s is much less than the natural lifetime of the excited state, the expected photon count given by (3) is saturated, that is, does not increase with the pulse energy. Fortunately, E and τ_L can both be chosen to optimize system performance. If we let $\tau_s = \alpha \tau_n$ and then solve (3) and (4) for the pulse energy and pulse length, we obtain

$$E = \left(1 + \frac{1}{\alpha}\right) \frac{16 z_s^2 h c N}{\eta T_A^2 \lambda \sigma_{\text{eff}} C_s r_0^2} \quad (5)$$

$$\tau_L = (1 + \alpha) \frac{21.5 t_n N}{\pi \eta T_A \lambda^2 C_s} \quad (6)$$

Equations (5) and (6) allow us to calculate the pulse energy

and pulse length required to generate a laser guide star in the sodium layer that is bright enough to operate an adaptive telescope of diameter D at its diffraction limit within an atmosphere characterized by r_0 . Notice that the required pulse energy is dependent on the diameter of the seeing cell.

The effective backscatter cross section depends on the laser line width and wavelength. The maximum value, which is easily obtained with current laser technology, is approximately $9 \times 10^{-16} \text{ m}^2$ (FWHM line width $< 0.5 \text{ pm}$). If we include a safety margin and require $N = 100$, then by substituting reasonable values into equations (5) and (6) ($T_A = 0.7$, $\eta = 0.075$, $\sigma_{\text{eff}} = 9 \times 10^{-16} \text{ m}^2$, $C_s = 5 \times 10^{13} \text{ m}^{-2}$, $r_0 = 0.25 \text{ m}$) the laser requirements for $\alpha = 1$ are calculated to be 88 mJ per pulse and a 24- μs pulse length. If mirror corrections are applied every 10 ms, the required laser power at 100 pps is 8.8 W. These values can be readily achieved with current dye-laser technology.

In addition to the restrictions already discussed, it is necessary to eliminate the low-altitude Rayleigh scatter signal from the wavefront sensor. Consequently, the quadrant detectors must be gated and the laser pulse length limited to about 300 μs . But by using a laser with a high repetition rate, the signals from several pulses can be averaged to improve the measurement accuracy. If photon noise is assumed to establish the fundamental performance limit of a laser-guided adaptive imaging system, then depending on the values of C_s and r_0 , 20–30 m is the largest telescope diameter that can be adaptively compensated for by using a guide star generated in the sodium layer.

The Mauna Kea experiments were conducted to demonstrate the feasibility of the technique by using a large astronomical telescope to photograph a laser guide star created in the sodium layer and to verify the predicted return flux. The University of Illinois (UIUC) lidar system was installed in an observatory dome with the laser mounted adjacent to a 0.6-m telescope. The collimated laser beam was projected toward the zenith with its

Table 1 Lidar and imaging system parameters

Lidar system	Imaging system
Laser: flashlamp-pumped dye	Telescope diameter: 2.2 m
Wavelength: 589 nm	Field of view: 1.7 mrad
Energy: 20 mJ/pulse	Optical bandwidth: 0.61 nm FWHM
Line width: 10.1 pm FWHM	Detector: 500 × 500 TI CCD
Pulse width: 2 μs FWHM	Quantum efficiency: 76%
Repetition rate: 7.5 pulse s^{-1}	Read-out noise: 15 counts
Beam divergence: 0.40 mrad FW @ e^{-2}	Latitude: 19°50' N
Telescope diameter: 0.6 m	Longitude: 155°28' W
Field-of-view: 0.4 mrad	Altitude: 4.2 km MSL
Optical bandwidth: 5 nm FWHM	
Receiver range gate: 150 m	
PMT: RCA C31034A	
Quantum efficiency: 15%	

own separate optical system. The lidar receiving subsystem was installed on the 0.6-m telescope and the complete lidar system was used to measure sodium density profiles during the imaging experiments. Table 1 summarizes the lidar and imaging system parameters.

The laser guide spot was imaged by a second telescope, the University of Hawaii 2.2-m telescope, located 117 m south-east of the laser. The separation between laser and receiver was large enough to prevent low-altitude Rayleigh scattered light from contaminating the laser spot that was formed in the sodium layer at an altitude of ~ 95 km. Reimaging optics followed by a 500×500 TI CCD (charge-coupled device) camera⁸ were placed at the Cassegrain focus of the 2.2-m telescope. A narrow-band interference filter was installed in the collimated beam following the reimaging optics.

By pointing all telescopes at the zenith, CCD images of the laser spot and lidar profiles of sodium density were recorded simultaneously. Figure 1 is an 8-min exposure of the laser-guide spot obtained starting at 13:43 UT on 21 January 1987. The lidar system was operated continuously during the exposure with a laser pulse rate of 7.5 pulse s^{-1} . Figure 2 is an expanded contour plot of the image photon count and Fig. 3 is the lidar profile of the sodium density measured during the exposure. Detailed parameters for each data set are given in the figure captions.

The laser spot shown in Figs 1 and 2 extends over a fairly large solid angle because the beam is not perfectly collimated. The spot dimensions are approximately $0.40 \text{ mrad} \times 0.58 \text{ mrad}$ ($36 \text{ m} \times 53 \text{ m}$) full width at the e^{-2} intensity points. The asymmetry is caused by the separation between the laser and the imaging telescope and by the finite thickness of the sodium layer. The angular width of the image is the smallest in the direction normal to a line joining the laser and telescope and is equal to the laser divergence angle.

The beam divergence must be reduced by a factor of ~ 100 to produce a guide star that can be used with a full adaptive imaging system. This problem is not significant for two reasons. First, lasers with better beam quality than the one used in this experiment are available today. Second, the full aperture of the adaptive telescope can, and for other reasons must, be used to transmit the laser beam³. The beam diameter for this experiment was only about 10 cm. Because the angular divergence of a collimated beam decreases as the beam diameter increases, with proper beam expansion the size of the laser guide spot can be reduced to the required value.

One of the primary goals of the Mauna Kea experiments was to calibrate the apparent brightness of the laser-guide star in terms of the observed sodium abundance. Because the laser did not saturate the sodium layer, a linear relation existed between the laser energy, sodium abundance, and the return flux N_p ,

$$N_p \approx T_A^2 \frac{\lambda ER \sigma_{\text{eff}} C_s}{4\pi (z_s - z_0)^2 hc} \quad (7)$$

where $R = 7.5$ is the laser pulse rate, $z_0 = 4.2$ km is the observatory altitude and the other variables are defined following (3). The UIUC laser line width is larger than the width of the sodium D_2 resonance line so that $\sigma_{\text{eff}} \approx 1.84 \times 10^{-16} \text{ m}^2$. The expected flux, calculated from (7) and the data in Table 1 and Figs 2 and 3, is approximately $4.03 T_A^2 \text{ s}^{-1} \text{ cm}^{-2}$. On the two nights when the laser spot was imaged with the CCD, we observed standard stars from the lists of Oke⁹ and Stone¹⁰ to calibrate the image brightness. The measured flux at the telescope was approximately $1.92 T_A \text{ s}^{-1} \text{ cm}^{-2}$, which implies that $T_A \approx 0.48$. This value for T_A is consistent with the value inferred from the Rayleigh scatter signal level at 30-km altitude measured by the lidar system. It is smaller than the values typically observed at Mauna Kea. However on this night, the low-altitude lidar profiles clearly show the presence of a pronounced 1-km thick dust layer just above the observatory and another pronounced aerosol layer extending from about 23 km to about 30 km. If a natural star produced this same photon flux when observed through a filter

with a 100-nm bandwidth centred at the visible wavelengths, it would have an apparent magnitude of $M_v = 14.4$.

Results from both the laser experiment and the analytic calculations confirm to first order that the laser-guide-star concept is viable and that moderately bright artificial guide stars can be created. Although much work needs to be done before a large diffraction-limited, ground-based telescope is put into operation, the potential scientific rewards clearly warrant the effort.

We thank Daniel Senft and Kevin Kwon for the painstaking work of installing the UIUC lidar system and making the laser operate at the 4.2 km altitude of the Mauna Kea Observatory; Bruce Barnes for providing valuable logistical support during the experiments, and the University of Hawaii for providing telescope time at the Mauna Kea Observatory; and Byron Welsh for his assistance in processing the laser-guide star images. Supported in part by NSF grants.

Received 6 March; accepted 8 May 1987.

1. Hardy, J. W. *Proc. Inst. elect. Electron. Engrs* **66**, 651-697 (1978).
2. Pearson, J. E., Freeman, R. H. & Reynolds, H. C. Jr in *Applied Optics and Optical Engineering* Vol. 7 (eds Shannon, R. & Wyant, J.) 245-340 (Academic, New York, 1979).
3. Foy, R. & Labeyrie, A. *Astr. Astrophys.* **152**, L29-L31 (1985).
4. Fried, D. L. *J. opt. Soc. Am.* **56**, 1372-1379 (1966).
5. Thompson, L. A. & Ryerson, H. H. *Proc. SPIE Conf. Instr. Astron.* **V**, 455, 560-568 (1984).
6. Gardner, C. S., Voelz, D. G., Sechrist, C. F. Jr & Segal, A. C. *J. geophys. Res.* **91**, 13659-13673 (1986).
7. Megie, G., Bos, F., Blamont, J. E. & Chanin, M. L. *Planet. Space Sci.* **26**, 27-35 (1977).
8. Hlivak, R. J., Pilcher, C. B., Howell, R. R., Colucci, A. J. & Henry, J. P. *Proc. SPIE Conf. Instr. Astr. IV*, 331, 96-103 (1982).
9. Oke, J. B. *Astrophys. J. Suppl. Ser.* **27**, 21-35 (1974).
10. Stone, R. P. S. *Astrophys. J.* **218**, 767-769 (1977).
11. Gardner, C. S. & Voelz, D. G. *J. geophys. Res.* **92**, 4673-4694 (1987).

Observations of the source and propagation of atmospheric gravity waves

G. Crowley* & P. J. S. Williams†

* Physics Department, Leicester University, Leicester LE1 7RH, UK

† Physics Department, University College of Wales, Aberystwyth SY23 3BZ, UK

Atmospheric gravity waves play an important role in transporting energy from one region of the atmosphere to another. But the sources of these waves and their propagation characteristics are not fully understood. Here we report on measurements made during the Worldwide Atmospheric Gravity Wave Study (WAGS) in the auroral source region showing that enhanced Joule heating and Lorentz forcing lead to the generation of large-scale thermospheric waves. Both of these source mechanisms depend on the auroral electric field and the intrinsic periodicity of that field determines the periodicity of waves observed an hour later propagating southwards over the UK. These observations provide the first confirmation of a link between a periodic auroral source and a large-scale gravity wave.

Magnetospheric processes drive large-scale currents through the high-latitude ionosphere at altitudes between 100 and 150 km. These currents cause Joule heating and in the same region the precipitation of energetic particles produces further atmospheric heating. The energy input in both cases can be calculated^{1,2} and during geomagnetically active intervals, especially at night, Joule and particle heating can exceed solar extreme ultraviolet and ultraviolet heating. Thus, the thermosphere in these regions is subject to strong thermal forcing to which must be added the Lorentz force produced by the interaction of auroral currents and the geomagnetic field. One of the effects of such forcing is thought to be the generation of atmospheric gravity waves (AGWs)³.

† Present address: HAO/NCAR, Boulder, Colorado.

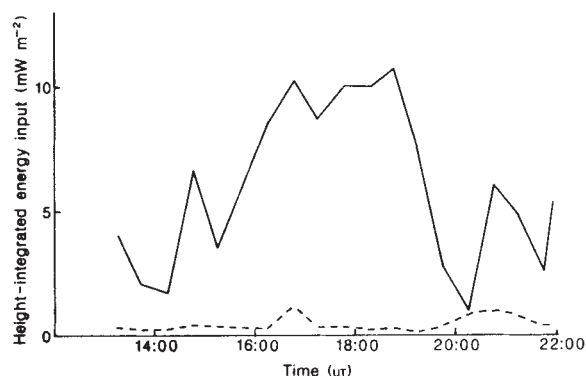


Fig. 1 Height-integrated heat input from Joule heating and particle heating on 18 October 1985. Solid line, Joule heating; dashed line, particle heating.

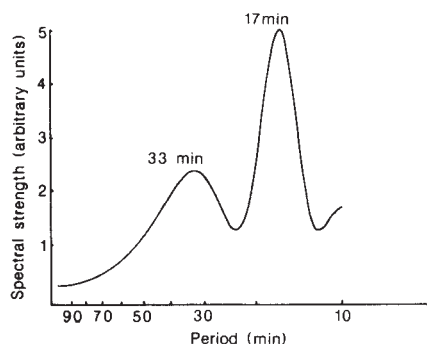


Fig. 2 Power spectrum of E^2 measured by EISCAT between 16:00 and 19:00 UT on 18 October 1985 (T. S. Viri, personal communication).

Theoretical work⁴ indicates that the mid-latitude response to an impulsive auroral event is quasi-periodic but consists of only one or two cycles. However, long-period waves ($\tau \sim 40$ min) continuing for many cycles have been reported⁵ and such sustained waves suggest that the source mechanisms are themselves periodic.

The WAGS campaign⁶ used a global network of instruments to make a comprehensive study of AGWs. In the European sector, auroral processes were observed by the European incoherent scatter radar (EISCAT)⁷, while ionosondes, HF-Doppler radar, meteor radar and a radio-telescope, were deployed to monitor wave propagation over the UK.

The EISCAT measurements used in the present paper were made above Tromsø (Norway) on 18 October 1985, providing electron concentration, electron and ion temperatures and plasma velocity over the height range 80–500 km. From the electron concentration the Pedersen conductivity of the ionosphere, σ_P , was determined while the component of plasma velocity perpendicular to the magnetic field provided measurements of the electric field E . From σ_P and E^2 Joule heating was calculated as a function of time. The profile of electron concentration also indicated the level of ionization due to energetic particles and from this the heating due to particle precipitation was derived. On average the height-integrated Joule heating was more than an order of magnitude larger than the particle heating (Fig. 1). Maximum Joule heating was observed at a height of 130 km between 16:00 and 19:30 UT, with a typical height-integrated value of 5–10 mW m⁻². After the passage of the Harang discontinuity at 20:00 UT, the increased hardness of precipitation in the westward electrojet lowered the height of maximum Joule heating to below 125 km, and the height-integrated average was reduced to ~ 3 mW m⁻², but this was still much larger than the average particle heating.

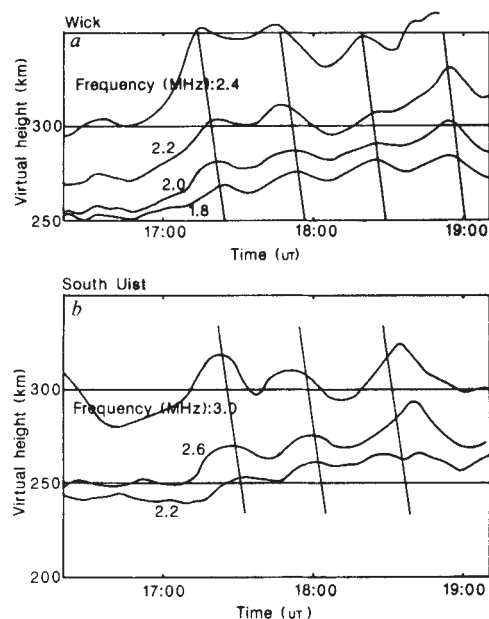


Fig. 3 Ionosonde measurements at *a*, Wick and *b*, South Uist on 18 October 1985 showing the descending wave-fronts of an AGW of period 33 min (G. Watkins, personal communication).

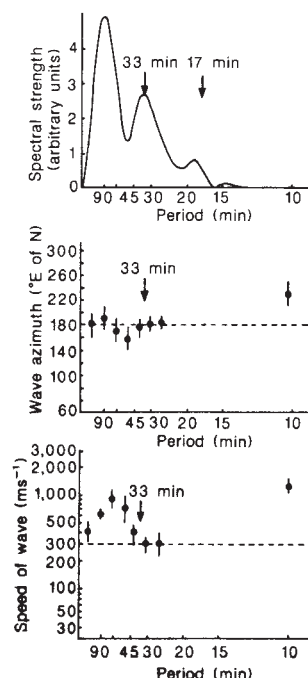


Fig. 4 Power spectrum of the wave observed by the Leicester HF-Doppler radar between 17:00 and 20:00 UT on 18 October 1985. The direction and speed of the waves were derived by cross-spectral analysis (I. McCrea, personal communication).

The relative importance of the Lorentz and Joule mechanisms (L/J) in contributing to long-period gravity waves at mid-latitudes depends critically on the height of the source, the electric field strength and the horizontal velocity of the wave⁸. At 130 km and above, for $E > 25$ mV m⁻¹, L/J is less than unity for all large-scale waves. As the electric field measured between 16:00 and 19:30 UT was about 30 mV m⁻¹ it is concluded that on this occasion Lorentz forces played a smaller part than Joule heating in launching the waves which were first observed over the UK about an hour afterwards.

The role of magnetospheric electric fields in generating AGWs was further investigated by comparing the periodic variations

in the electric field with the periods of the waves observed. A power-spectrum analysis of both E and E^2 was performed using a three-hour window. Two peaks were observed corresponding to periods of 33 and 17 min (Fig. 2). It is likely that the 33-min period represented the characteristic time for storage and release of energy in the magnetosphere.

The same periods were present in the waves observed over the UK. Between 17:00 and 19:00 UT the descending wavefronts characteristic of AGWs were detected by ionosondes at Wick and South Uist, with periods in the range 32 to 35 min (Fig. 3). Waves were also measured by an HF-Doppler network centred at Leicester⁹. In addition to the usual long-period component, which is thought to be the mid-latitude response to an impulsive source, the spectra of these all showed a clear peak at ~33 min with a much weaker peak at ~17 min. Assuming that these waves were generated in the auroral zone, the small amplitude of the shorter-period wave at mid-latitudes was consistent with its preferential attenuation, as predicted by theory¹⁰. Cross-spectral analysis of the data taken by the network indicated that the 33-min wave propagated towards the equator at a speed of ~350 m s⁻¹ (Fig. 4).

The onset of enhanced Joule heating and Lorentz forcing in the eastward electrojet shortly before 16:00 UT was caused by

an electric field with intrinsic periodicities of 33 and 17 min. About an hour later waves with similar periodicity were observed propagating southwards over the United Kingdom. These observations provide direct evidence for an auroral source of AGWs.

We acknowledge the remarkable effort made by the whole WAGS community in the United Kingdom and Dr K. Schlegel (Max-Planck Institut für Aeronomie, Lindau) in planning and carrying out a very ambitious programme of coordinated measurements and in reducing the data. Special thanks are due to the Director and Staff at EISCAT. EISCAT is an international facility supported by France, Finland, West Germany, Norway, Sweden and the United Kingdom.

Received 9 March; accepted 30 April 1987.

1. Banks, P. M. *J. atmos. terr. Phys.* **39**, 179–193 (1977).
2. Brekke, A. & Rino, C. L. *J. geophys. Res.* **83**, 2517–2524 (1978).
3. Hunsucker, R. D. *Rev. Geophys. Space Phys.*, **20**, 293–315 (1982).
4. Testud, J. thesis, Univ. Paris (1973).
5. van Eyken, A. P., Williams, P. J. S., Maude, A. D. & Majid, A. J. *J. atmos. terr. Phys.* **42**, 513–515 (1980).
6. Argo, P. & Hunsucker, R. D. *Eos* **12 March**, 121–122 (1985).
7. Folkestad, K., Hagfors, T. & Westerlund, S. *Radio Sci.* **18**, 867–879 (1983).
8. Brekke, A. J. *J. atmos. terr. Phys.* **41**, 475–479 (1979).
9. Waldoock, J. A. & Jones, T. B. *J. atmos. terr. Phys.* **48**, 245–260 (1986).
10. Richmond, A. D. *J. geophys. Res.* **84**, 1880–1890 (1979).

Thermoelectric power and electron–phonon enhancement in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$

A. Mawdsley*, H. J. Trodahl*, J. Tallon‡, J. Sarfati† & A. B. Kaiser§

* Departments of Physics, and † Chemistry, Victoria University of Wellington, Private Bag, Wellington, New Zealand

‡ Physics and Engineering Laboratory, Department of Scientific and Industrial Research, Private Bag, Lower Hutt, New Zealand

§ Max-Planck-Institut für Festkörperforschung, D-7000 Stuttgart 80, FRG

The recent discovery of high-temperature oxide superconductors^{1–4} has generated enormous interest both among physicists and, because of their obvious technological importance, in the wider community. One of the less-studied but potentially interesting properties is the thermoelectric power (or thermopower), which is the coefficient that characterizes the electromotive force resulting from the redistribution of electrons in a temperature gradient. The temperature dependence of the thermopower is modified by the electron–phonon interaction, and provides a means of estimating the strength of the interaction. We have performed measurements on several samples of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$, and find an unexpected precursor effect just above the transition temperature T_c —namely, an increase rather than decrease in thermopower. The measured thermopower is positive with a magnitude typical of metallic conductors, and shows some similarity to the calculated diffusion thermopower for a relatively large electron–phonon coupling parameter of $\lambda \approx 2.5$.

The samples were prepared by mixing the oxides and/or carbonates of yttrium, barium and copper in the appropriate proportions, reacting in air and sintering in oxygen. The resulting material has been inspected by electron microscopy and X-ray diffractometry to establish that the concentration of the $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ superconducting phase is >95%. The oxygen decrement $\delta = -0.05$ was determined by thermogravimetric analysis in hydrogen, and the sample densities are ~80% of the theoretical value. The samples show a relatively narrow transition into the superconducting state (<0.5 K for 0–50%), with resistance falling to zero at ~91–93 K. Above the transition

temperature the resistivities rise by ~20% in the first 20 K, and thereafter they show a more gradual rise, with the resistivity approximately doubling between 110 K and room temperature. The absolute value of the room-temperature resistivity varies among the samples, with samples containing the most dense packing of the pure $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ phase showing the lowest resistivity. No resistivity anomaly at 240 K has been seen in any of these samples.

Thermoelectric power measurements on the disk-shaped samples were made in a cryostat that has been used extensively in our studies of amorphous metal and semiconducting thin films. Temperatures were measured with gallium arsenide light-emitting-diode sensors and the thermoelectric voltage was measured with manganin leads which have been calibrated against lead.

Figure 1 shows the resistivity and thermoelectric power of one of our most conductive samples. As in all of our measurements, the thermopower is zero in the superconducting state (as required by thermodynamic arguments). Immediately above T_c the thermopower (S) falls sharply from $4 \mu\text{V K}^{-1}$, but above 110 K it rises slowly and linearly with temperature. The initial fall is clearly related to the accompanying rapid rise in the resistivity (ρ) between 91 and 110 K, and we suspect that both are precursors to the superconducting transition. The character of the precursor is rather surprising, because one would expect the precursor effect in thermopower, like that in resistivity, to be a tendency towards its zero value in the superconducting state. Superconducting pairs are able to counteract a net diffusion of electrons between regions of different temperature without the need for an applied voltage, so the effect of superconductive pairing should be to decrease the thermoelectric voltage. That this does not occur suggests that incipient pair formation decreases the normal-state conductivity more strongly above the Fermi energy than below, leading to a positive contribution to the diffusion thermopower.

The thermopower is positive across the entire (normal-phase) temperature range. The phonon-drag thermopower is normally suppressed in metals with resistivities above $100 \mu\Omega \text{ cm}$ (ref. 5), so we interpret our results as a diffusion thermopower, which can be written:

$$S = X_b T [1 + \lambda^* \lambda_s(T)]$$

where $X_b = S_b/T$ is the bare thermopower parameter, λ^* is the effective electron–phonon enhancement at low temperatures, and $\lambda_s(T)$ gives the decay of thermopower enhancement with

† Permanent address: Victoria University of Wellington, Department of Physics, Private Bag, Wellington, New Zealand.

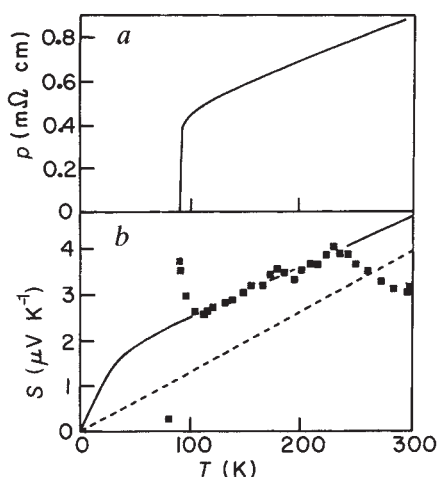


Fig. 1 Resistivity (a) and thermoelectric power (b) of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ plotted against temperature. The solid line in b is the fit to the data using the Eliashberg function of Weber, and the dashed line is the assumed bare thermopower.

temperature⁶. The renormalization of electron energy by the electron-phonon interaction leads to the thermopower slope at low temperatures being $(1+\lambda^*)$ times the slope of the bare thermopower S_b . There are some additional effects, involving velocity and relaxation-time renormalization⁶ and higher-order diagrams⁷, that change the effective enhancement value of λ^* , but these appear to be small in glassy metals and are neglected. The form of λ_s is determined by the Eliashberg function (α^2F) where α represents the strength of the electron-phonon interaction and F is the phonon density of states and in Fig. 1 we compare our data to the enhancement based on α^2F for the La-Ba-Cu oxides⁸, which have a similar phonon spectrum to $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. The electron-phonon enhancement effect in Fig. 1 is calculated without adjustable parameters, the absolute magnitude of the thermopower being determined by the multiplication factor X_b , which is chosen to give similar absolute values for the calculation and the data.

It can be seen that the effect of the electron-phonon enhancement is to change the thermopower from linear behaviour going through the origin (assumed for the bare thermopower S_b) to behaviour that is still approximately linear over most of the temperature range but not extrapolating to the origin. Such behaviour is seen in high-resistivity Chevrel-phase compounds¹⁰ (in which phonon drag is suppressed), and excellent agreement between the calculated and observed values of S is obtained. The Y-Ba-Cu-O data appear to follow this pattern between 100 K and 230 K.

The value of $\lambda^* \approx 2.5$ estimated from the thermopower is very much larger than those in common metals^{11,12}, consistent with the very strong electron-phonon coupling required for high-temperature superconductivity with the conventional phonon-mediated pairing mechanism. It should be noted that the recently reported absence¹³ of an isotopic shift of T_c argues against any phonon-mediated mechanism¹⁴, but there remains the curious coincidence of an abnormally high electron-phonon enhancement with a high T_c .

The magnitude of the bare thermopower parameter is given by the Mott formula¹⁵:

$$X_b = (\pi^2 k^2 / 3e) [d(\ln \sigma(\epsilon)) / d\epsilon],$$

where σ is the contribution of electrons with energy ϵ to the electrical conductivity, and k and e are the Boltzmann constant and electronic charge. The derivative is to be evaluated at the Fermi energy ϵ_f . The energy dependence of the conductivity results primarily from the electron density of states and the relaxation time, so that the energy

$$\epsilon_0 = \pi^2 k^2 / 3e X_b$$

gives the energy scale over which these change. (For a free-electron gas, $\epsilon_0 \approx \epsilon_f$). The values of X_b and ϵ_0 are within the range commonly found in metals; in fact, these values are nearly identical to those of pure copper¹⁶. It is interesting that the electron density of states (per unit volume) is also very close to those of elemental metals, rather than the somewhat larger values found in 20-K superconductors such as Nb_3Ge (ref. 17).

Finally, we interpret the fall in the thermopower above 250 K as the result of a competing high-temperature conduction path through the (green) semiconducting compound Y_2BaCuO_5 , known to coexist in small concentrations with the superconducting compound. The thermopower contribution from these paths will be weighted by their (small) contribution to the electrical conductance, but because the semiconducting thermopower is much larger than that of a metal, significant deviations result. This argument is supported by measurements on high-resistivity samples, which show the negative departure from metallic behaviour at a lower temperature.

We thank Albert Schuitema for technical assistance and Ken McKenzie and Peter Gilberd for X-ray diffractometry and electron microscopy. A.B.K. acknowledges the support of an Alexander von Humboldt Fellowship.

Received 5 June; accepted 23 June 1987.

1. Bednorz, J. G. & Müller, K. A. *Z. Phys.* **B64**, 189-193 (1986).
2. Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908-910 (1987).
3. Cava, R. J. *et al. Phys. Rev. Lett.* **58**, 1676-1679 (1987).
4. Murphy, D. W. *et al. Phys. Rev. Lett.* **58**, 1888-1890 (1987).
5. Jäckle, J. *J. Phys. F*, **10**, L43-L46 (1980).
6. Kaiser, A. B. *Phys. Rev. B*, **29**, 7088-7091 (1984).
7. Kaiser, A. B. & Stedman, G. E. *Solid St. Commun.* **54**, 91-94 (1985).
8. Weber, W. *Phys. Rev. Lett.* **58**, 1371-1374 (1987).
9. Renker, B. *et al. Z. Phys. B*, (in the press).
10. Kaiser, A. B. *Phys. Rev. B*, **35**, 4677-4681 (1987).
11. Bergmann, G. *Phys. Rep.* **27**, 159-185 (1976).
12. Grimvall, G. *The Electron-Phonon Interaction in Metals* (North-Holland, Amsterdam, 1981).
13. Batlogg, B. *et al. Phys. Rev. Lett.* **58**, 2333 (1987).
14. Anderson, P. W. & Abrahams, E. *Nature* **327**, 363 (1987).
15. Blatt, F. J., Schroeder, P. A., Foiles, C. L. & Greig, D. *Thermoelectric Power of Metals* (Plenum, New York, 1976).
16. MacDonald, D. K. C. *Thermoelectricity: An Introduction to the Principles* (Wiley, New York, 1962).
17. Batlogg, B., Ramirez, A. P., Cava, R. J., van Dover, R. B. & Rietman, E. A. *Phys. Rev. B*, **35**, 5340-5342 (1987).

Fast macromolecules control mutual diffusion in polymer blends

Russell J. Composto*, Edward J. Kramer* & Dwain M. White†

* Department of Materials Science and Engineering and the Materials Science Center, Cornell University, Ithaca, New York 14853, USA

† General Electric Corporate Research and Development, Schenectady, New York 12301, USA

The tracer diffusion coefficient D^* of small concentrations of tagged linear polymers seems well understood on the basis of the reptation model¹⁻⁶, but the mutual diffusion coefficient D which characterizes diffusion at higher concentrations remains controversial⁷⁻¹⁷. Such diffusion is important in blends of miscible, yet chemically dissimilar, polymers where it controls the kinetics of phase separation and the welding of polymer interfaces. The controversy centres on predicting D from the D^* s of the components of the blend. While the 'slow theory'^{7,8} predicts that D is controlled by the diffusion of the slower-moving component, 'fast theory'^{9,10} says it is controlled by the diffusion of the faster-moving one. Here we report experiments in which we measure the mutual diffusion coefficient as we increase the tracer diffusion coefficient of the faster-moving component in the blend by decreasing its molecular weight. The strong increase in D we observe demonstrates that the 'fast theory' prediction is correct.

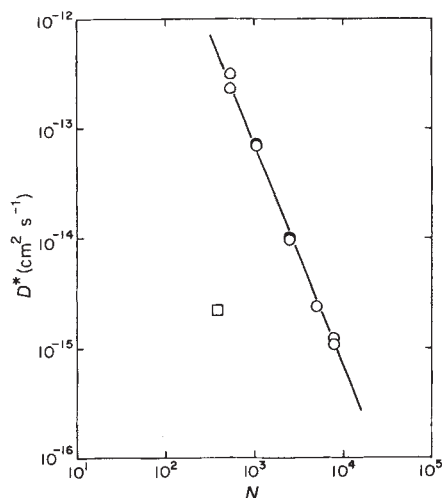


Fig. 1 Tracer diffusion coefficients D^* of d-PS (circles) and d-PXE (squares) as a function of the weight average degree of polymerization N of the d-polymers. The diffusion was at 206 °C into a protonated blend matrix of 0.55 volume fraction PS ($M_w = 2,000,000$) and PXE ($M_w = 35,000$). The solid line is a line of slope -2 as predicted by reptation.

Using the Flory-Huggins theory¹⁸ the mutual diffusion coefficient D can be written as⁷⁻¹⁰

$$D = 2(\chi_s - \chi)\phi_A\phi_B D_T \quad (1)$$

where χ is the Flory segment-segment interaction parameter and χ_s is its value at the spinodal and $\phi_A\phi_B D_T$ is an Onsager transport coefficient where ϕ_A and ϕ_B are the volume fractions of polymers A and B. The 'slow theory'^{7,8} requires that the fluxes of the polymers cancel so that

$$D_T^{-1} = \phi_B(D_A^* N_A)^{-1} + \phi_A(D_B^* N_B)^{-1} \quad (2)$$

where N_A and N_B are the degrees of polymerization of the A and B polymers respectively. The 'fast theory'^{9,10} requires that osmotic pressure gradients vanish so that

$$D_T = \phi_B D_A^* N_A + \phi_A D_B^* N_B \quad (3)$$

Thus D is controlled by the diffusion of the fastest-moving species for the 'fast theory' and by the slowest-moving species for the 'slow theory'.

To provide a critical test for the theories we measured both the D^* s and D for a miscible blend of polystyrene(PS):poly(2,6-dimethyl-1,4-phenylene oxide) also named poly(xylenyl ether) (PXE). The D^* s were found from the concentration versus depth profiles, measured by forward recoil spectrometry (FRES)¹⁹, of trace amounts of the deuterated polymers d-PS and d-PXE that were diffused into a blend of the protonated polymers which had a volume fraction $\phi = 0.55$ of PS. Bilayer films of a very thin layer of the pure deuterated polymer (~ 20 nm thick) on top of a thick layer of the blend were diffused in vacuum ($< 10^{-6}$ torr) at 206 °C and the FRES spectra were analysed by methods described previously^{5,14,19}. The d-PS or d-PXE from the thin layer becomes rapidly diluted so that the combinatorial entropy of mixing is the only driving force for diffusion over most of the diffusion time and thus a true tracer diffusion coefficient is measured. The d-PXE was synthesized at General Electric²⁰ and had a weight average molecular weight of $M_w = 46,000$ ($N_{PXE} = 383$) and a polydispersity index PDI of 2.4. The d-PS were nearly monodisperse polymers (PDI < 1.3) with M_w s ranging from 53,000 to 820,000 ($N_{PS} = 510$ to $N_{PS} = 7,885$) purchased from Polymer Laboratories. The data on polydisperse polymers were analysed to yield the D^* corresponding to the weight average molecular weight M_w ²¹. The M_w and PDI of the protonated PXE and PS in the blend matrix were 35,000 and 2.3 and 2,000,000 and < 1.2 respectively. These values were high

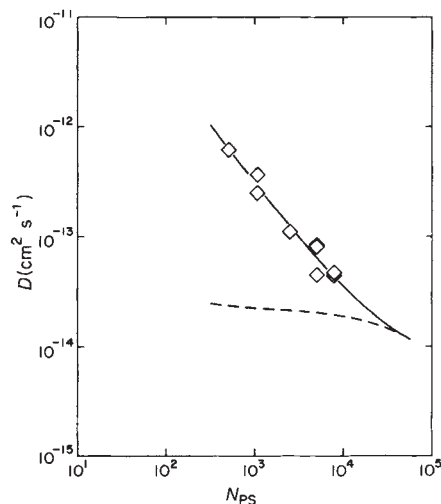


Fig. 2 The mutual diffusion coefficient in the d-PS:PXE blend system at 206 °C and a 0.55 volume fraction of d-PS as a function of the degree of polymerization of d-PS; the degree of polymerization of PXE = 292 for all samples. The solid line is the prediction of the 'fast theory' whereas the dashed line is the prediction of the 'slow theory'.

enough to ensure that reptation was the predominant mechanism of diffusion of both the d-PS and d-PXE. The D^* s of the d-PS and d-PXE are plotted as a function of N of the tracer polymer in Fig. 1. Note that D^* of d-PS scales as N^{-2} as predicted by reptation¹⁻⁶. Note also that D^* of PXE is nearly two orders of magnitude smaller than D^* of PS at the same value of N .

The mutual diffusion couples consisted of two films of deuterated polystyrene (d-PS):protonated PXE blends which have a small difference in the volume fraction ϕ of d-PS, from $\phi = 0.50$ in the ~ 350 nm thick top film to $\phi = 0.60$ in the ~ 2 μ m-thick bottom film¹⁴. The initial step function d-PS concentration profile was broadened by mutual diffusion at 206 °C in vacuum ($< 10^{-6}$ torr). The volume fraction versus depth profile of d-PS was measured by FRES and a mutual diffusion coefficient D was extracted by fitting these profiles to a standard error-function solution¹⁴; the D measured corresponds to the D at the average composition of the diffusion couple, $\phi = 0.55$.

Figure 2 shows that D increases as the degree of polymerization N_{PS} of the d-PS chains decreases, scaling approximately as N_{PS}^{-1} . Since PS is the fastest-moving species in these blends the observed variation of D with N_{PS} constitutes strong qualitative support for the 'fast theory'. An even stronger quantitative test of the two theories can be performed by computing D from the measured D^* . (The D^* s for the d-PS were used directly; the D^* for the 35,000 PXE was found by scaling the D^* for the 46,000 d-PXE by the N^{-2} expected for reptation.) In the computation χ_s is set equal to its mean field value $0.5[(\phi N_{PS})^{-1} + ((1-\phi)N_{PXE})^{-1}]$ and χ is taken as -0.0178 as determined previously¹⁴; the $(\chi_s - \chi)$ term in equation (1) is insensitive to N_{PS} . The solid line in Fig. 2 shows the D computed using the 'fast theory', that is, equations (1) and (3) whereas the dashed line shows the D computed using the 'slow theory', that is, equations (1) and (2). Clearly the data are in excellent agreement with the 'fast theory' and in strong disagreement with the 'slow' theory. Thus we have directly demonstrated for the first time that the mutual diffusion of a polymer blend is controlled by the diffusion of its faster-moving species.

Brochard and de Gennes¹⁶ have proposed a hybrid, 'fast-slow theory' that predicts for diffusion distances less than

$$l = (D_A^* \tau_B)^{1/2} = \langle R_B^2 \rangle^{1/2} (D_A^* / 3\pi^2 D_B^*)^{1/2} \quad (4)$$

that the slower-moving B macromolecules are swollen like a gel by the faster-moving A ones so that the 'fast theory' result holds,

whereas for greater diffusion distances an osmotic pressure gradient develops so that the 'slow theory' result obtains. In equation (4) τ_B is the reptation time and $\langle R_B^2 \rangle^{1/2}$ is the root mean square end-to-end distance of the B chains. This idea seemed to explain the results of marker displacement experiments^{9,12} in which the slower B molecules were extremely long (20,000,000 M_w PS) since the l_s , which ranged from 1 μm to 20 μm , were larger than the diffusion distances that could be probed by the ion-beam technique used. In the present experiments, however, the slower PXE molecules are rather short ($\langle R_B^2 \rangle^{1/2} = 16 \text{ nm}$) so that the l_s computed for these PS:PXE blends range from 11 to 39 nm, much smaller than the typical diffusion distances (Dt)^{1/2} $\approx 200 \text{ nm}$ measured. The Brochard and de Gennes theory¹⁶ therefore requires that our D follow the 'slow theory' prediction, equation (2), with which our experimental results strongly disagree. Recent measurements of concentration profiles of deuterated and protonated polybutadienes which have different N_s and diffusion distances $>100 \mu\text{m}$ suggest similar conclusions²².

This work was supported by the Division of Materials Research, NSF Polymers program, and we benefitted from the use of the facilities of the Cornell Materials Science Center which is funded by the Division of Materials Research-Materials Research Laboratory programme of the NSF. We especially thank J. W. Mayer for his encouragement and appreciate useful

discussions and correspondence with H. Sillescu, K. Binder, J. H. Wendorff, G. Fytas, P. G. de Gennes, J. Klein, R. A. L. Jones and A. M. Donald.

Received 26 February; accepted 8 May 1987.

1. de Gennes, P. G. *J. chem. Phys.* **55**, 572-579 (1971).
2. Doi, M. & Edwards, S. F. *JCS Faraday Trans. II* **74**, 1789-1801 (1978).
3. Klein, J. *Nature* **271**, 143-145 (1978).
4. Antonietti, M., Coutandin, J., Gruetter, R. & Sillescu, H. *Macromolecules* **17**, 798-802 (1984).
5. Green, P. F., Mills, P. J., Palmström, C. J., Mayer, J. W. & Kramer, E. J. *Phys. Rev. Lett.* **53**, 2145-2147 (1984).
6. Tirrell, M. *Rubber Chem. Tech.* **57**, 523-556 (1984).
7. Brochard, F., Jouffroy, J. & Levinson, P. *Macromolecules* **16**, 1638-1641 (1983).
8. Binder, K. *J. chem. Phys.* **79**, 6387-6409 (1983).
9. Kramer, E. J., Green, P. & Palmström, C. J. *Polymer* **25**, 473-480 (1984).
10. Sillescu, H. *Makromol. Chem. rapid Commun.* **5**, 519-523 (1984).
11. Gilmore, P. T., Falabella, R. & Lawrence, R. L. *Macromolecules* **13**, 880-883 (1980).
12. Green, P. F., Palmström, C. J., Mayer, J. W. & Kramer, E. J. *Macromolecules* **18**, 501-507 (1985).
13. Jones, R. A. L., Klein, J. & Donald, A. M. *Nature* **321**, 161-162 (1986).
14. Composto, R. J., Mayer, J. W., Kramer, E. J. & White, D. M. *Phys. Rev. Lett.* **57**, 1312-1315 (1986).
15. Marshall, U., Fischer, E. W., Herkt-Maetzky, Ch. & Fytas, G. *J. Polymer Sci. Lett.* **24**, 191-197 (1986).
16. Brochard, F. & de Gennes, P. G. *Europhys. Lett.* **1**, 221-224 (1986).
17. Garaballa, R. W. & Wendorff, J. H. *Macromolecules* (in the press).
18. Flory, P. J. *Principles of Polymer Chemistry* 507-511 (Cornell University Press, Ithaca, 1953).
19. Mills, P. J., Green, P. F., Palmström, C. J., Mayer, J. W. & Kramer, E. J. *Appl. Phys. Lett.* **45**, 957-959 (1984).
20. White, D. M. & Nye, S. A. *Macromolecules* **17**, 2643-2644 (1984).
21. Mills, P. J., Green, P. F., Palmström, C. J., Mayer, J. W. & Kramer, E. J. *J. Polymer Sci.* **24**, 1-9 (1986).
22. Jordan, E. A. *et al. Macromolecules* (in the press).

Anomalous elastic thickness of the oceanic lithosphere in the south-central Pacific

Stéphane Calmant & Anny Cazenave

CNES/GRGS, 18 Avenue Edouard Belin, 31055 Toulouse Cedex, France

To first order the elastic thickness T_e of the oceanic lithosphere under mid-plate volcanoes increases linearly with the square root of plate age at the time of loading¹⁻⁴. A large scatter around this simple relationship is observed however, suggesting that factors other than age influence the elastic thickness and produce departure from the first-order trend⁵. To look at second-order effects, we have conducted a worldwide analysis over the Pacific, Atlantic and Indian oceans and determined T_e under sixty volcanoes, applying elastic flexure theory and using as observational constraints, geoid height data from the Seasat satellite. We find that for the three main oceans, the depth of the elastic layer under volcanoes increases with age of plate at loading time t , according to the simple relationship $T_e = 2.70 \pm 0.15 t^{1/2}$ (T_e in km, t in Myr), except under all of the volcanic chains of the south-central Pacific where this relation does not apply. The elastic layer is found to be much thinner there, and indicates a broad regional anomaly, presumably of thermal origin, well correlated with several other geophysical anomalies reported in this region⁶⁻¹⁰.

Volcanic chains and isolated seamounts supported by elastic flexure of lithospheric plates produce large positive geoid anomalies up to 10-m amplitude and 100-500-km wavelength. These anomalies result from two opposite effects: (1) a positive geoid anomaly due to the topography and (2) a negative anomaly due to flexure of the plate under the load. The flexural response of the lithosphere, modelled using the equilibrium theory of deformation of thin elastic loaded plates overlying a weaker fluid medium, is characterized by its flexural rigidity D or equivalently its elastic thickness T_e . Comparing observations (for example, geoid height), with theoretical estimates allows us to calculate a best fit of D and hence T_e . Using the synthetic bathymetric profiling systems (SYNBAPS) bathymetric data¹¹ given over a 5×5 arc min grid to compute the topographic load

and geoid height data from the Seasat satellite as model constraints, we have applied a three-dimensional numerical method in the spatial domain to compute (i) the flexural deformation of the lithosphere under the load and (ii) the geoid anomaly from which D is estimated (see refs 12-14 for details on the computation method). We assumed a crustal thickness of 6 km, a topography and crust density of $2,800 \text{ kg m}^{-3}$, and a mantle density of $3,400 \text{ kg m}^{-3}$. We also assumed that the material filling the deflection has the same density as the topography itself. For each volcano, the theoretical geoid has been computed for several values of D over a regular grid covering an area several times larger than the volcano extent. Comparison between the computed and observed geoid is then performed by sampling the three-dimensional computed geoid along all Seasat altimeter profiles (long wavelengths $\geq 4,000 \text{ km}$ removed) crossing the volcano in the immediate vicinity of its summit. This procedure which ensures $\sim 7\text{-km}$ along-track resolution in geoid height is preferred to a direct comparison with the gridded Seasat geoid as this describes the short-wavelength anomalies due to the topography less accurately, a result of the large ($\sim 50 \text{ km}$) spacing between Seasat tracks. A minimization of the sum of the residual squares criterion is applied along each satellite profile to select the best-fitting D value. The associated standard error in D is evaluated from the r.m.s. curves. Uncertainties in D estimates are from three different contributions: (i) uncertainties in the bathymetry; (ii) uncertainties in the assumed densities (mainly densities of the volcanoes and of the material filling the deflection) and (iii) the adopted flexure model of isostasy (for example, continuous versus fractured elastic plate model). Seasat data, on the other hand with $\sim 10 \text{ cm}$ accuracy in the geoid height, do not in general introduce significant errors. From the above three contributions, errors in the SYNBAPS data induce the largest discrepancies between the computed and observed geoid. Nevertheless, cases of mislocation of the load are detected by the non-coincidence of topography and geoid maxima above the volcano. Using $2,800 \text{ kg m}^{-3}$ for the topography density may overestimate the load. But tests using $2,700 \text{ kg m}^{-3}$ have shown that changes in the estimates of D remain, in general, within the error bars. Adoption of a fractured elastic plate instead of a continuous plate model would have led to systematically higher estimates of D , but relative variations from one site to another presumably would remain the same. The selection of volcanoes considered in this study depends on

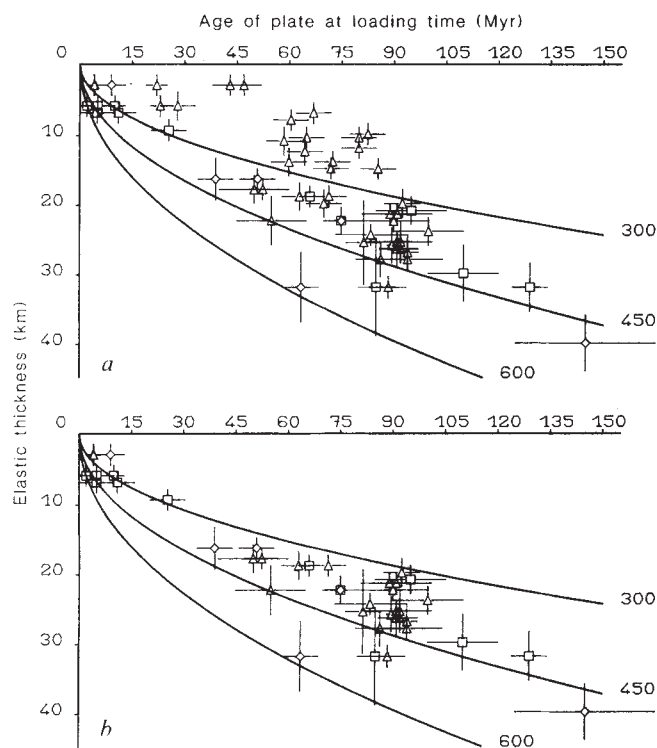


Fig. 1 *a*, Elastic thickness T_e estimates under sixty volcanoes distributed in the three main oceans as a function of age of plate at the time of loading. Symbols: triangles refer to the Pacific volcanoes, squares refer to the Atlantic volcanoes and diamonds to the Indian ocean volcanoes. Solid curves represent the isotherms of the cooling half-space model. *b*, Same as Fig. 1*a* but excluding the south-central Pacific (Marquesas, Pitcairn, Society and Cook-Austral chains) T_e values.

the availability of reliable age estimates (the age of the crust and the age of the load) as well as having closely spaced satellite tracks giving a good description of the geoid anomaly above the volcano. For long volcanic chains, D has been determined for separate volcanoes along the chain. To estimate the age of the crust at the sites of the studied volcanoes, we used the isochron map of Larson *et al.*¹⁵ and linearly interpolated along successive isochrons. We used published dating for the ages of the volcanoes. The elastic thickness T_e was deduced from D assuming E (Young's modulus) = 10^{11} N m⁻² and ν (the Poisson ratio) = 0.25.

We have determined T_e at forty-one volcanoes distributed along linear chains (Hawaiian, Emperor, Marquesas, Pitcairn, Society, Easter and Cook-Austral chains in the Pacific ocean) and at nineteen individual volcanoes: Samoa and Valerie Guyot (Louisville Ridge) in the Pacific, Bermuda, the Azores, Corner and Atlantis (New England seamounts), Madeira, Cruiser, Great Meteor, Cape Verde Rise, Ascension, St Helene in the Atlantic Ocean, Réunion, Crozet, Comores, Prince Edward, Rodriguez and Mauritius Island in the Indian Ocean. The Pacific Ocean results (location of volcanoes, age of crust, age of load and T_e estimates) are presented in ref. 14. The Atlantic and Indian Ocean results (not yet published) are available on request. The elastic thickness T_e versus the age of plate at the time of loading t is plotted in Fig. 1*a* where no clear trend emerges at first glance. A more detailed inspection, however, reveals two distinct groups: (1) volcanoes of the south-central Pacific (Cook-Austral, Society, Pitcairn and Marquesas chains) are associated with a very low elastic thickness compared to loads emplaced on plates of similar age in other regions; (2) all remaining observations confirm results of previous studies, that is, T_e increases with the square root of lithospheric age at the

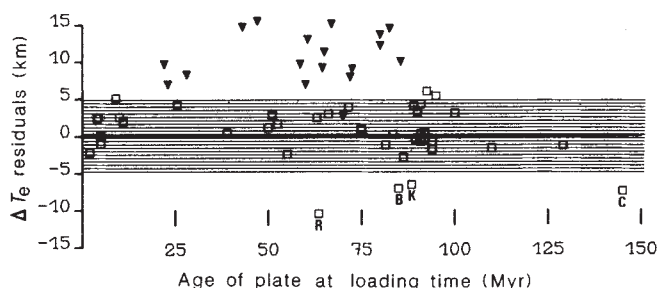


Fig. 2 Plot of ΔT_e residuals: average trend ($T_e = 2.70 t^{1/2}$) minus T_e estimates versus age of plate at the time of loading. Symbols: B, Bermuda; C, Comores; K, Kauai (Hawaiian chain); R, Reunion Island. Black triangles at more than 5 km above the average trend refer to the south-central Pacific volcanoes.

time of loading with very good agreement from one ocean to another as shown by Fig. 1*b*, identical to Fig. 1*a* except for the south-central Pacific results which have been excluded. Applying a linear regression of the data shown in Fig. 1*b*, we find that the effective elastic thickness of the oceanic lithosphere under volcanoes follows the approximate relationship

$$T_e = 2.70 \pm 0.15 t^{1/2} \quad (1)$$

(T_e is in km if t is in Myr). This relationship suggests that the depth of the elastic layer under the volcanoes follows the depth of the $400 \pm 20^\circ\text{C}$ isotherm of the cooling half-space model¹⁶. Figure 2 shows a plot of ΔT_e residuals (average trend as given by T_e calculated from equation (1) minus observed T_e). All points more than 5 km above the average trend correspond to the south-central Pacific volcanoes. Four points (Reunion, Bermuda, Kauai in the Hawaiian Islands, Comores) have a negative ΔT_e (observed T_e is larger than the average trend). These cases correspond to large emerged islands for which the topography may have been underestimated which leads to an overestimate of T_e .

The low- T_e values found for the south-central Pacific indicate a thinning of the elastic lithosphere over a broad region. The largest discrepancies are observed for the eleven volcanoes of the Cook-Austral chain (average thinning of the elastic layer amounting to 13 km) while ~ 8 -km thinning is observed at Society, Marquesas and Pitcairn seamounts. An exception is Tahiti Island for which a normal elastic thickness is observed. The low elastic thickness observed for the south-central Pacific volcanoes affects a broad area east of the East Pacific Rise on the Pacific plate, extending from $\sim 10^\circ\text{S}$ to $\sim 30^\circ\text{S}$ and from 120°W to $\sim 160^\circ\text{W}$. More to the west, Samoa Island (14.2°S , 169.5°W) has a normal elastic thickness (24 km for $t = 100$ Myr). This is also the case for a seamount of the Louisville Ridge located to the south-west (41.5°S , 164.3°W) where $T_e = 18$ km for $t = 50$ Myr. The south-central Pacific low elastic thickness is correlated with several other geophysical anomalies observed in that region: (1) high concentration of hotspots; (2) unusually shallow seafloor depth⁶ compared to the normal subsidence predicted by thermal cooling models¹⁶; (3) anomalously low subsidence rate of the ocean floor⁷; (4) anomalous surface wave velocities⁸; (5) anomalously thin thermal plate derived from geoid step data across fracture zones⁹ and (6) the broad pattern of lineated gravity undulations of ~ 10 mgal amplitude and ~ 200 km wavelength, discovered in the Seasat altimeter data¹⁰.

McNutt and Fischer⁶ have identified a broad area of extended depth anomalies, covering French Polynesia and the Cook-Austral chain. They call this area the south Pacific superswell. They favour a thermal origin for the superswell and show that depth anomalies are consistent with a thin lithosphere of thermal thickness ~ 70 km, compared to ~ 125 km in the north Pacific and the north Atlantic¹⁶. From bathymetric profiles, anomalously low subsidence rates have been reported by Coch-

ran⁷ on the western flank (Pacific Plate) of the East Pacific Rise, between 9° S and 22° S and ascribed to a hotter underlying asthenosphere. A broad slow velocity anomaly of surface-wave velocities has been reported by Nishimura and Forsyth⁸ over the south-central Pacific. This anomaly centred at ~20° S, 135° W covers the Marquesas, Society, Cook-Austral, Pitcairn and Easter hotspots and is associated with shallow seafloor depth and low elastic thickness. The analysis of Seasat-derived geoid height data at several fracture zones of the south Pacific led Cazenave⁹ to invoke an apparently thinned thermal plate (~70 km) for the young lithosphere. But these observations which indicate a more complex lithospheric thermal structure at fracture zones than predicted by conductive cooling models, extends farther to the south (up to ~55° S) than those cited above (low elastic thickness, superswell, low subsidence rates, slow surface wave velocities). These independently observed anomalies call for a common origin associated with a complex thermal structure of the lithosphere in the south-central Pacific and possibly of the underlying asthenosphere. Indeed, the Seasat-derived gravity lineations discovered by Haxby and Weissel¹⁰ between the East Pacific Rise and French Polynesia are ascribed by these authors to small-scale convective instabilities developing in a thin low-viscosity zone beneath the lithosphere, as predicted by convection models with depth-dependent vis-

cosities^{17,18}. Small-scale convection enhanced by low viscosity at the base of the plate is also invoked to explain the superswell⁶ as well as geoid anomalies at south Pacific fracture zones^{18,19}. The low values of the elastic plate thickness found here under the volcanoes of the south-central Pacific suggest that the effects of the anomalous thermal structure of the asthenosphere extend up to shallow depths inside the lithosphere itself.

Received 22 December 1986; accepted 5 May 1987.

1. Watts, A. B. *J. geophys. Res.* **83**, 5985-6004 (1978).
2. Watts, A. B. *J. geophys. Res.* **84**, 3817-3826 (1979).
3. Watts, A. B., Bodine, J. H. & Ribe, N. M. *Nature* **283**, 532-537 (1980).
4. Watts, A. B., Karner, G. D. & Steckler, M. S. *Phil. Trans. R. Soc. A* **305**, 249-281 (1982).
5. McNutt, M. *J. geophys. Res.* **89**, 11180-11194 (1984).
6. McNutt, M. & Fischer, K. *Eos* **67**, 1220 (1986).
7. Cochran, J. R. *Geophys. J. R. astr. Soc.* **87**, 421-454 (1986).
8. Nishimura, C. E. & Forsyth, D. W. *Geophys. J. R. astr. Soc.* **81**, 389-407 (1985).
9. Cazenave, A. *Earth planet. Sci. Lett.* **70**, 395-406 (1984).
10. Haxby, W. F. & Weissel, J. K. *J. geophys. Res.* **91**, 3507-3520 (1986).
11. Van Wyckhouse, R. J. *SYNBAPS (Synthetic Bathymetric Profiling Systems) Technical Report* (Naval Oceanographic Office, Washington DC, 1973).
12. Cazenave, A. & Dominh, K. *J. geophys. Res.* **89**, 11171-11179 (1984).
13. Calmant, S. & Cazenave, A. *Earth planet. Sci. Lett.* **77**, 187-202 (1986).
14. Calmant, S. *Earth planet. Sci. Lett.* (in the press).
15. Larson, R. L. *et al. The Bedrock Geology of the World* (Freeman, New York, 1985).
16. Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803-827 (1977).
17. Buck, W. R. *Nature* **313**, 775-777 (1985).
18. Buck, W. R. & Parmentier, E. M. *J. geophys. Res.* **91**, 1961-1974 (1986).
19. Craig, C. H. & McKenzie, D. *Earth planet. Sci. Lett.* **78**, 420-426 (1986).

Evidence for two zones of debris entrainment beneath the Greenland ice sheet

D. E. Sugden*, P. G. Knight†, N. Livesey‡, R. D. Lorrain§, R. A. Souchez§, J.-L. Tison¶ & J. Jouzel¶

* Department of Geography, University of Edinburgh, Drummond Street, Edinburgh, EH8 9XP, UK

† Department of Geography, University of Aberdeen, AB9 2UF, UK

‡ Department of Soil Science, University of Aberdeen, AB9 2UF, UK

§ Laboratoire de Géomorphologie, Université Libre de Bruxelles, B1050, Belgium

¶ Laboratoire de Géochimie Isotopique, LODYC/DPC/CEN Saclay, France

Investigation of the ice, debris and stable isotope characteristics of basal ice exposed at the margin of the Greenland ice sheet provides a means of testing and refining models of ice-sheet erosion and deposition. The basal ice exposed at the ice-sheet margin gives evidence of the processes operating at the bed at different locations, which can be used to constrain theoretical models predicting varying basal conditions^{1,2}. Improved understanding of the effects of isotope fractionation in hydrogen and oxygen caused by freezing processes at the bed which modify the isotope characteristics of the basal ice layers enables us to report evidence of two separate zones of debris entrainment beneath the ice sheet.

Greenland is the only place where there is an extensive terrestrial ice-sheet margin where the crucial basal ice sequence can be examined *in situ*. Data were collected from two sites, one at the snout of the Russell Glacier near Søndre Strømfjord representing a lobate margin, and one at the northern flank of the Jakobshavn Isbrae outlet glacier, representing a marginal zone of streaming flow (Fig. 1). Flow lines at the Jakobshavn site suggest a local origin for the ice studied, whereas at Russell Glacier some of the ice is of a more distant, inland origin.

Some glaciological characteristics of the two sites, including ice morphology, type, structures and, in the case of Russell Glacier, approximate thickness are shown in Fig. 2. The ice temperatures have not yet been measured at either site, but approximate mean annual temperatures can be estimated by



Fig. 1 The location of the Russell Glacier and Jakobshavn Isbrae in West Greenland.

altitudinal extrapolation from the nearby settlements of Søndre Strømfjord (-5.1 °C) and Jakobshavn (-5.7 °C) to be ~-7 °C at Russell Glacier and ~-6.5 °C at Jakobshavn Isbrae. Melt water can be seen flowing from beneath Russell Glacier in summer suggesting that part of the glacier base is at the pressure melting point at least seasonally. Ice flow is increasingly compressive towards the ice margin. Measurements along a 600-m stretch of ice margin of the Russell Glacier during a four-week period in September 1986, revealed horizontal velocities (30 m from the margin of the glacier) of 73-180 m yr⁻¹ and compressive strain rates of -0.39 to -2.0 yr⁻¹; within a few metres of the margin strain rates rose to -4.0 yr⁻¹. Ablation rates in September averaged 30-40 mm per day.

The stratigraphy of the two sites is shown in Fig. 2. The dominant feature in the extreme marginal zone in both cases is the presence of subparallel debris-rich ice layers up to 10 cm thick dipping steeply up glacier. The debris bands occur both as discrete layers of debris, with only interstitial ice, and as fine debris laminations, the two types often merging into each other laterally or vertically. Cobbles and boulders associated with the

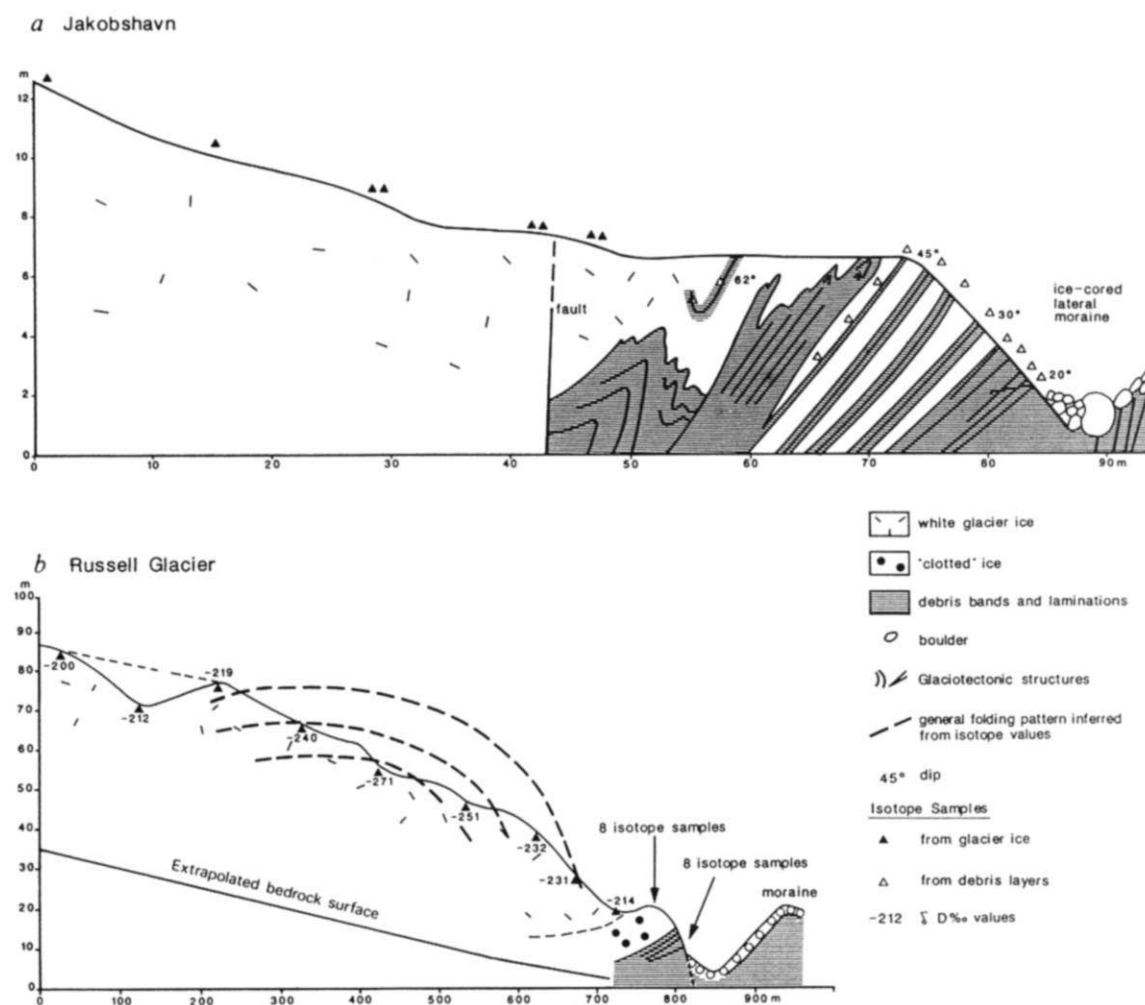


Fig. 2 The stratigraphy of the marginal ice at Jakobshavn Isbrae and the Russell Glacier, Søndre Strømfjord. *a*, The banded debris series at the northern margin of Jakobshavn Isbrae, showing folds and isotope sample sites. *b*, Profile of the Russell Glacier showing three ice types and $\delta^{18}O$ values. The latter show that $\delta^{18}O$ values become less negative up the glacier as would be expected in terms of normal glacier flow, but variations demonstrate the existence of a large discontinuity or fold. The approximate bedrock surface has been extrapolated from adjacent ice-free slopes.

sequence are generally aligned along debris bands but may truncate them. Clearly visible ice structures show the whole sequence is substantially folded, with the limbs of the folds more or less perpendicular to the ice flow. The ice intimately associated with the debris bands (type A), exists as lenses up to 1 cm thick within bands and as ice laminae within the laminated debris. It is clear, bubble free, and often in layers of only one crystal thickness.

At Jakobshavn the marginal zone in which type A occurs is directly overlain by white, bubbly, virtually clean glacier ice (type C), with crystals of 1–1.5 cm.

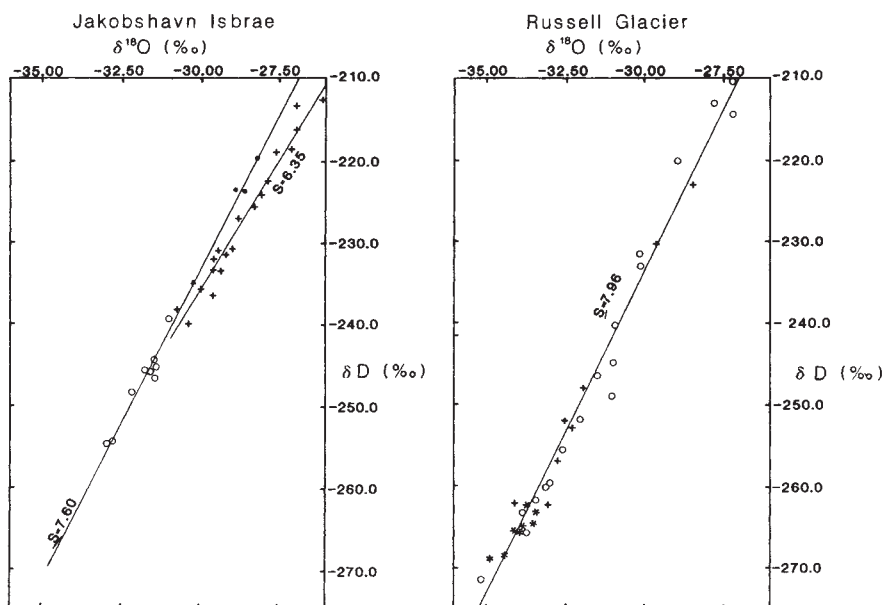
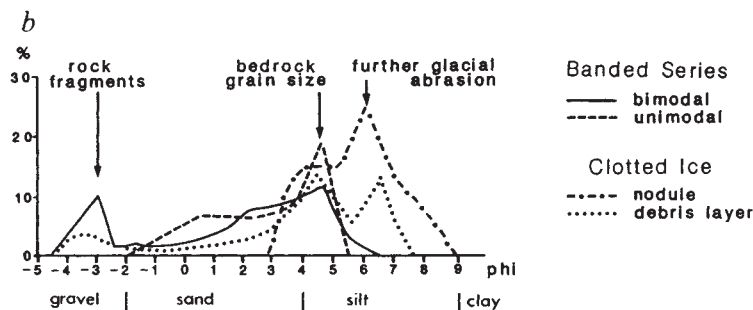
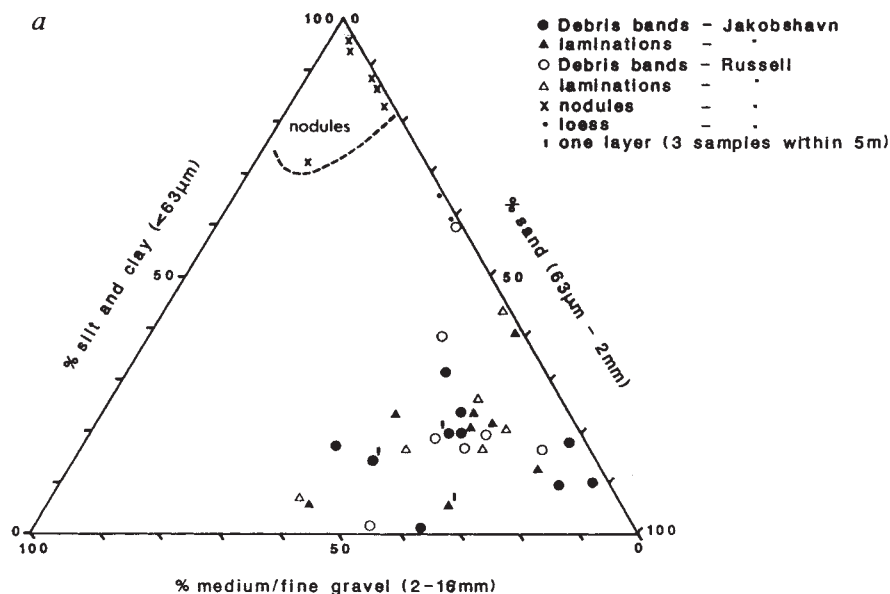
At the Russell Glacier site a third type (type B) occurs in a layer up to ~10 m thick between the debris band sequence and the clean glacier ice. This ice, which we refer to as 'clotted ice' (type B), was not observed at the Jakobshavn site. The ice is essentially clear with crystals of 2–5 cm, the crystal boundaries sometimes marked by pockets of gas. It contains suspended lenticular nodules of debris up to 8 cm in diameter, the size and concentration of which decrease towards the top of the sequence (Table 1). In places the nodules occur in planar clusters associated with small bubbles and occasional angular clasts.

The debris in ice of type A is mainly sand and fine gravel, with considerable variation in the size distribution of individual samples, even from a single band (Fig. 3a). Eighty per cent of samples show a clear grain size frequency peak in the coarse

silt range (0.23–0.06 mm, ~4 phi) and in 34% of the samples there is a second peak at 6–20 mm (~–3 phi, Fig. 3b). This bimodal distribution is typical of glacial crushing³, with one peak representing the grain size of the gneissic parent bedrock and the other rock fragments. The lithology of the sediment mirrors the local bedrock. None of the samples from the debris bands contains clay-size particles typical of glacial abrasion. Probably this reflects the periodic flushing of the ice-rock system by melt water, illustrated by the presence of clays in melt-water streams and marginal lakes in the area.

Table 1 Debris contents in different types of basal ice, Russell Glacier and Jakobshavn Isbrae

	Sample number	Mean debris content by % weight
Banded series (type A)		
Debris bands Jakobshavn	10	55
Debris bands Russell	3	71
Debris laminations Jakobshavn	8	34
Debris laminations Russell	6	14
Clotted ice (type B)		
Nodule (7 × 5.5 × 4 cm)	1	76
Minimum (upper layers)	2	0.001
Maximum (lower layers)	3	4–8



The debris in the nodules in ice of type B is different from that in the banded sequence in one major respect, namely that it contains a finer component with 71-97% of the weight comprising silt and clay-sized particles ($<63\ \mu\text{m}$) (Figs 3a and 3b). Silt is dominant, but the clay-sized fraction varies from 0-8.5% from nodule to nodule. The fine debris is consistent with an origin from gneissic bedrock, as illustrated by the clay mineral assemblage (Table 2).

Co-isotope analysis of δD and $\delta^{18}\text{O}$ in ice can distinguish unaltered glacier ice from ice which has undergone refreezing

Fig. 3 Grain size of the rock debris. *a*, Triangular diagram showing the coarse nature and variability of debris in the banded series and the finer component in the nodules of the clotted ice. Within the banded series there is considerable variation in three samples taken within a 5-m length along one debris band but no difference between debris in laminations and that in compact bands. *b*, The peaks in coarse silt and gravel fractions are typical of processes of subglacial crushing. The finer component due to abrasion is missing from the banded series but present in the nodules and debris concentrations in the 'clotted' ice of Russell Glacier.

Fig. 4 The δD - $\delta^{18}\text{O}$ diagrams for samples from Jakobshavn Isbrae (left) and Russell Glacier (right). Open circles represent glacier ice; crosses, type A ice; asterisks, type B ice and black dots lake water. The slopes of the different regression lines are indicated by *S*. Sampling procedure and sample analysis are as in ref. 6.

and determine the isotopic composition of the parent water involved⁴⁻⁶. For this study 66 ice and water samples from different locations at both sites were analysed for their isotopic composition and the results are shown in Fig. 4.

The $\delta\text{D}/\delta^{18}\text{O}$ values from glacier ice (type C) at the top of the sequence at both sites lie along a regression line close to the value of eight normal for unaltered precipitation⁷. At Jakobshavn these glacier ice samples are within a narrow range ($\sim 2\%$ in $\delta^{18}\text{O}$) which is consistent with a local origin of all the ice, whereas at Russell Glacier the range of values ($\sim 8\%$ in

Table 2 Constituent minerals of the clay-sized material in 3 clotted (type B) ice samples and their relative importance

Mica-trioct	3	4	4
Chlorite	2	3	2
Amphibole	3	2	2
Quartz	2	2	2
Pyroxene	2	2	2
Feldspar (plagioclase)	2	2	3
Feldspar (alkali)	1	1	1
1.0–1.4 nm interstratified	2	—	1

(1, trace; 2, <10%; 3, 10–40%; 4, >40%). The assemblage is typical of an origin from gneissic bedrock.

$\delta^{18}\text{O}$) is even wider than the climatically induced isotope difference between modern and Pleistocene ice in the Camp Century core. This indicates the presence of ice from a range of sources on the ice sheet and probably a Pleistocene age for the most negative values.

Isotope samples from the debris bands (type A) at Jakobshavn lie along a slope of 6.3, demonstrating that the debris-bearing ice has been accreted at the bed by freezing on, a process first described by Weertman⁸. Samples from marginal lakes are less negative than the intersection point between the precipitation slope and the freezing slope. This discounts the hypothesis that marginal lakes supply the water admitted to freeze at the bed, since the parent water for the freezing process must lie at this intersection point⁴. The clarity of the freezing slope for the Jakobshavn samples is due to the isotopic homogeneity of the parent water supplied for freezing by the glacier ice above. The range of values between the least negative sample on the freezing slope and the parent water is ~4‰ in $\delta^{18}\text{O}$, too great to be accounted for by a single freezing event, which can only produce isotope enrichment of up to 3‰. This implies that ice in the debris layers has undergone multiple melting and freezing events. In contrast, type A samples from the Russell Glacier are not connected by a single freezing slope. Such a situation has already been described⁶, but we have no decisive argument here to consider mixing of water from different origins in the course of freezing.

The samples from ice of type B (clotted ice) cannot be isotopically distinguished from the most negative glacier ice samples at the site. This eliminates the possibility that the ice was accreted by bulk, large-scale, freezing on of water at the bed. If one accepts that the debris and bubble-free characteristics of this ice reflect basal freezing, then such a situation is likely to be associated with small-scale regelation as the ice passes over bumps. In this case water is melted upstream of a bump in very small amounts. Each incremental input of this water refreezes completely on the downstream side. Fractionation occurs only within individual frozen increments too small to be picked up at the scale of our sampling. Although the clotted ice is much thicker (≤ 10 m) than can be explained by bump-related regelation alone, plastic flow of basal ice around bed obstacles can significantly increase the thickness of the debris-bearing basal layer⁹. The very negative values of the regelation ice reflect the occurrence of Pleistocene glacier ice at the bottom since fractionation effects have not been detected.

The existence of these two zones of debris entrainment beneath the western part of the Greenland ice sheet, one near the margin and one further inland can be equated with zones of active glacial erosion beneath the ice sheet.

The 'marginal' zone produces the banded debris series and an origin by freezing on is demonstrated by the $\delta\text{D}/\delta^{18}\text{O}$ relationships, the low bubble density and the observation that individual ice layers are often just one crystal thick between dirt layers. Freezing on takes place in areas where melt water has previously flushed out fine abrasion products. The series is thickened by compressive flow near the ice margin, as demon-

strated by folds, faulting and measurements. Debris entrainment takes place near the ice margin as shown by the regularity and continuity of debris bands in an active tectonic environment, by the lithology and, in the case of Jakobshavn, by the isotope values of the parent water involved.

The 'interior' zone of entrainment produces the clotted ice. A longer distance of transport is suggested by the discontinuous nature of the debris inclusions, the fining upwards and the absolute values of δD and $\delta^{18}\text{O}$ which point to a parent source away from the ice margin. Accretion takes place in a closed melt water environment which entrains fine abrasion products, probably in association with pressure induced regelation. Such an interpretation is consistent with the presence of angular debris. The existence of several metres of this 'interior' regelation ice is significant for ice-core studies. It is not immediately distinguishable on isotope grounds from unaltered glacier ice. But the ice is recognizable in terms of its debris content, the nature of which resembles that described from several ice cores, for example, in Antarctica¹⁰, Greenland¹¹ and the Agassiz ice cap¹².

The field work was generously supported by The Royal Society and NATO. Maureen Lamb processed the sediment samples. P.G.K. was supported by a NERC studentship.

Received 17 December 1986; accepted 6 May 1987.

1. Budd, W., Janssen, D. & Radok, U. *Polarforschung* **6**, 39, 293–306 (1970).
2. Sugden, D. E. *Arctic Alpine Res.* **9**, 21–47 (1977).
3. Haldorsen, S. *Boras* **10**, 91–105 (1981).
4. Jouzel, J. & Souchez, R. A. *J. Glaciol.* **28**, 35–42 (1982).
5. Souchez, R. A. & Jouzel, J. *J. Glaciol.* **30**, 369–372 (1984).
6. Souchez, R. A. & De Groote, J.-M. *J. Glaciol.* **31**, 229–232 (1985).
7. Dansgaard, W. *Tellus* **16**, 436–468 (1964).
8. Weertman, J. *J. Glaciol.* **3**, 965–978 (1961).
9. Boulton, G. S. in *Ice Ages: Ancient and Modern* (eds Wright, H. E. & Moseley, F.) 7–42 (Seel House, Liverpool, 1975).
10. Gow, A. J., Epstein, S. & Sheely, W. J. *Glaciol.* **23**, 185–192 (1979).
11. Herron, S. & Langway, C. C. *J. Glaciol.* **23**, 193–208 (1979).
12. Gemmell, A. M. D., Sharp, M. J. & Sugden, D. E. *Earth Surf. Processes Landforms* **11**, 123–130 (1986).

Ice growth in supercooled solutions of antifreeze glycoprotein

K. Harrison*, J. Hallett*, T. S. Burcham†§, R. E. Feeney†, W. L. Kerr‡ & Y. Yeh‡

* Desert Research Institute, PO Box 60220, Reno, Nevada 89506, USA

† Department of Food Science and Technology, 1480 Chemistry Annex, University of California, Davis, California 95616, USA

‡ Department of Applied Science, University of California, Davis, California 95616, USA

Inhibition of ice growth in supercooled solution by certain proteins is vital to the survival of many living organisms. Some fish, native to both subzero northern and southern waters, have special proteins or glycoproteins in their blood serum that inhibit ice formation. Whereas these proteins have only a very small effect on the melting temperature of ice, the temperature of these fish can fall to nearly 1 K below the melting point before ice crystals grow^{1–3}. This phenomenon is called freezing hysteresis, in contrast to the normal colligative effect of solutes that depresses the equilibrium temperature, around which small changes lead to crystal growth or melting depending on sign. Some insects also exhibit a serum freezing hysteresis⁴. We report the effects of different degrees of supercooling on the habit and rates of growth of ice crystals from solutions of these antifreeze glycoproteins (AFGPs). We find that the crystallization rate is up to five times greater than that in pure water.

§ Present address: Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA.

Early work suggested that AFGPs might function at the ice-solution interface^{1,5}, and recent studies show an interfacial mechanism⁶⁻⁷. Further insight into the mechanism for inhibition of ice growth can be obtained from the habit of crystals that grow beyond the critical temperature for hysteresis. Raymond and DeVries⁸ reported that in solutions of AFGP, ice grew as long needles aligned with the ice *c*-axis. Recently it was shown that at very low concentrations of AFGP, ice crystals exhibited growth directions not seen before⁹. Unique forms of growth were also seen when the crystal growth front in an AFGP solution on a microscope slide developed stepwise and at right angles to the temperature gradient⁷. To investigate unequivocally the influence of the proteins on the habit of ice crystals, it is important that growth occurs in the absence of contact with solid boundaries, which may serve as nucleation sites or sinks of heat¹⁰.

To isolate the influence of different solutions and supercooling alone, a system was devised that nucleated crystals in the middle of a uniformly supercooled sample. Alternatively, single crystals of selected orientation were inserted into the free liquid surface.

In the first technique, 8 ml of solution was cooled to a fixed temperature, which was uniform within the sample to ± 0.05 K as measured by a thermocouple that traversed the volume. Nucleation was achieved by inserting a liquid-nitrogen-cooled, thin stainless-steel wire into the liquid in a 1-mm plastic tube that ended in the bulk of the liquid. In the second technique, an ice crystal of known orientation selected from the surface of a freezing dish of water was slowly inserted into the surface by micrometer control. For both the subsequent crystal growth was recorded by Video Cassette Recorder through a Questar telescope 2 m from the cell. Growth rates of the leading crystal tips were measured and the form of the crystals estimated visually. Observations in a Schlieren optical system indicate that local fluid convection was unimportant in these particular studies. The AFGP used was a mixture of the larger fractions, termed AFGP 1-5, isolated from the blood serum of *Pagothenia borchgrevinkii*¹¹. Ovomuroid, a glycoprotein used as a control, was isolated from chicken egg white.

Ice growth in AFGP solution is quite different from growth in other solutions. There is no detectable growth ($< 1 \mu\text{m s}^{-1}$) down to a specific temperature; after the ice crystal has been inserted into the solution it remains quite unchanged, neither melting nor growing. Once the freezing temperature of this solution is reached and the critical supercooling is surpassed, growth rates increase dramatically. For a range of further supercooling, the growth rate increases only slowly with supercooling but is significantly greater than the growth rates of pure water crystals (Fig. 1), and also greater than the enhanced growth rates found previously in 0.8 M NaCl solution¹². Beyond supercooling of 2 °C, the growth rate of ice in AFGP solution again increases with increasing supercooling but at a slower rate than found in the pure ice-water system. The growth rates given were measured from dendrites growing in the focal plane of the optical system, outward from the capillary tip; growth rates indicated in Fig. 1 are reproducible to an accuracy of 10% for individual crystals that remain in focus.

The growth habit of ice in AFGP solution behaves quite differently from pure water (Fig. 2). Initial growth beyond the critical supercooling is in the form of fine needles to be compared with narrow dendrites (Fig. 2*b, d* and *a, c*). At lower temperatures, in pure water, the external periphery of the dendrite arms tends to a hexagonal shape (2*e*) but for AFGP, it takes different, ill-defined directions (2*f*). As the supercooling approaches 2 °C, the growth forms revert next to that of pure water (2*g*), with the appearance of flat dendrites with well-defined facets parallel to the basal plane. The single crystal nucleation experiments show quite distinctly the orientation of the new crystals in relation to the nucleating crystal. In pure water (2*g*) a crystal with *c*-axis out of the page gives rise to a crystal growing as a thin dendrite in the plane of the photograph.

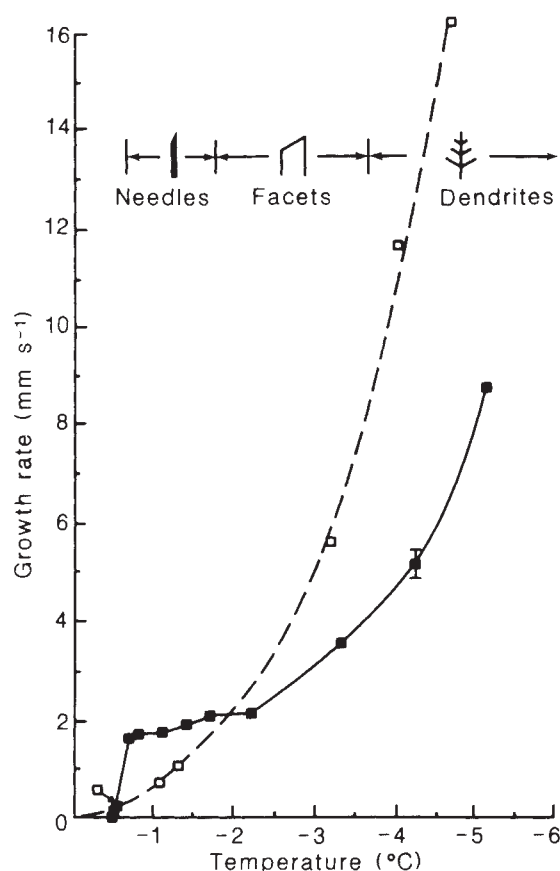


Fig. 1 Temperature dependence of growth rate of ice in pure water (□) and in 5 gl^{-1} solution of AFGP 1-5, (■). For this concentration of AFGP the polymer is functioning as an antifreeze agent until ~ 0.5 °C supercooling. Growth morphology for ice growing in AFGP solution is shown at the top of the figure.

By contrast, with a crystal orientated with *c*-axis perpendicular to this direction (in the plane of the page) crystals grow unequivocally parallel to the *c*-axis (2*h*). These results contrast with the control protein studies, where ovomucoid in solution leads to ice growth the same as that found in pure water.

In interpreting our results we must bear in mind the concept of how ice grows in pure water. At low supercooling (0.2–0.4 °C) ice grows in the form of thin disks. With increase of supercooling beyond 0.4 °C these disks develop an instability, and grow out as coarse dendrites. At further supercooling below –3 °C, the growth direction deviates progressively from the basal plane increasing to an angle of 10° at –5 °C. These observations are consistent with continuous growth on a molecularly rough surface in a direction perpendicular to the *c*-axis and on a molecularly smooth surface in the *c*-axis direction. As supercooling increases, two-dimensional nucleation (which requires a critical supercooling) takes place on this surface as well, leading eventually to continuous growth and to the observed change of growth direction.

We suggest that AFGP molecules succeed in blocking the rough growth perpendicular to the *c*-axis completely, as well as inhibiting surface nucleation necessary for *c*-axis growth on the basal plane, until the solution freezing temperature (critical supercooling point) is reached. This inhibition enables supercooling to be uniformly experienced over the entire surface of the ice, and eventually leads to nucleation and growth in the *c*-axis direction at temperatures just beyond the initial supercooling point. Here a thin needle grows, its tip favoured for thermal dissipation but its lateral growth still blocked by adsorbed molecules. With increasing supercooling, growth along the *a*-axis eventually reverts to its dominant role, and 'normal'

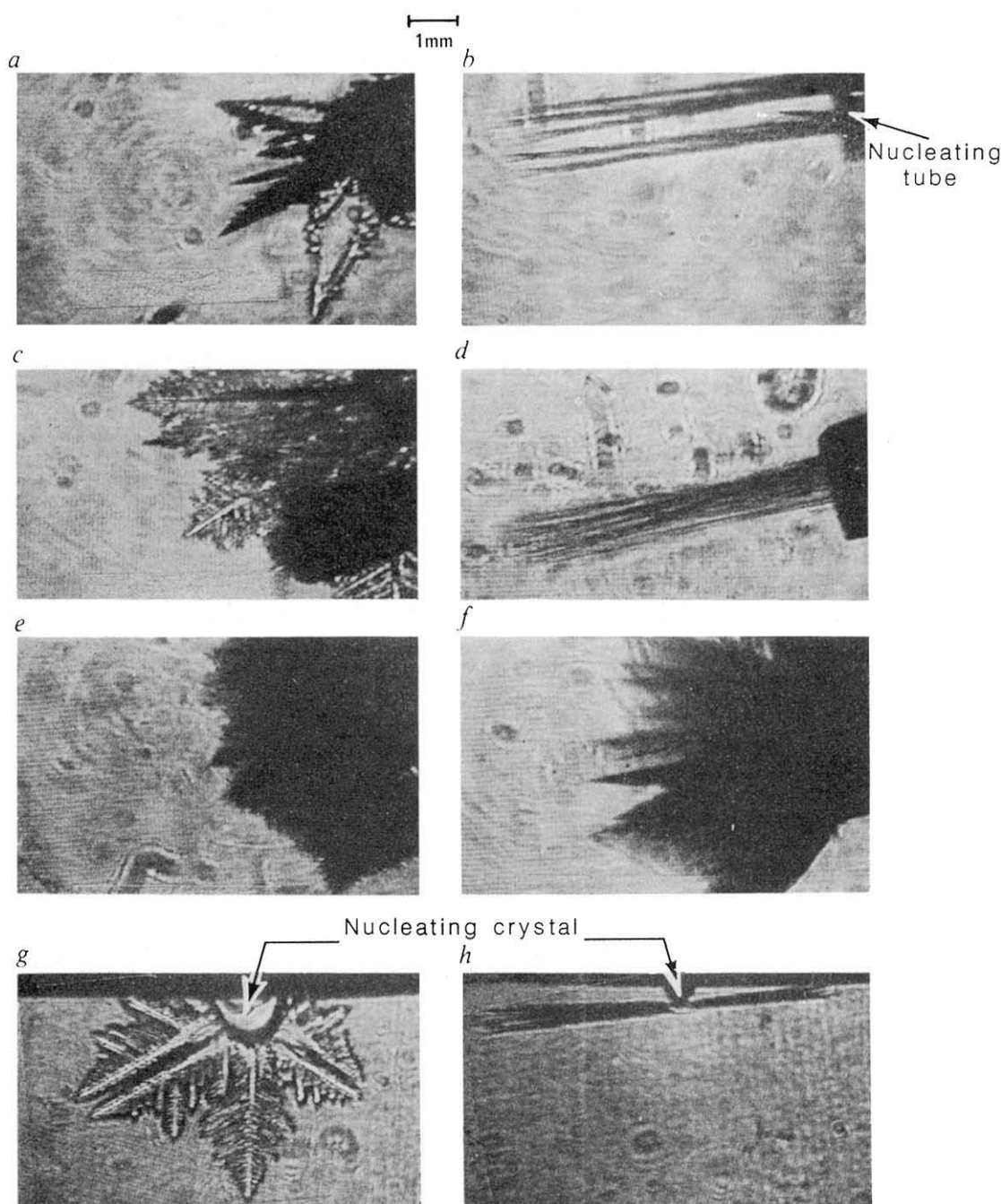


Fig. 2 Crystal growth of ice in pure water (*a, c, e, g*) and in 5 gl^{-1} AFGP solution (*b, d, f, h*) at nearly corresponding temperatures of supercooling in horizontal pairs. *a, b*, $T = -0.5^\circ\text{C}$; *c, d*, $T = -1.1^\circ\text{C}$; *e, f*, $T = -3.3^\circ\text{C}$; bottom pair (*g, h*), growth from single crystal of ice at $T = -1.25^\circ\text{C}$. In *g*, the *c*-axis is normal to the plane of the page; growth is along the *a*-axis. In *h*, the *a*-axis is normal to the plane of the page; growth is along the *c*-axis (see text).

growth behaviour continues. We argue that AFGP molecules serve to block growth only as long as the rate of addition of water molecules to an ice surface is below a rate (supercooling dependent) that allows the AFGP molecules time to adsorb and block water molecule adsorption onto the same site. Even though there apparently is a difference in these competitive rates along the two growth directions, the reason for this difference is not clear and requires investigation.

These results substantiate the structural observations of Knight *et al.*⁹ on the effects of AFGP molecules on ice growth. Most importantly, the present work shows that the rate of ice-crystal growth is very much altered when growth starts beyond the critical supercooling point. Implications of this phenomenon for freeze damage of cells cannot be overlooked.

Work on growth of crystals in a convecting environment funded by NASA; on growth of ice crystals from the vapour by NSF (J.H.). Also supported by NIH (R.E.F.) and NSF (Y.Y.).

Received 29 December 1986; accepted 6 April 1987.

1. Feeney, R. E. *Am. Scient.* **62**, 712-719 (1974).
2. DeVries, A. L. *Phil. Trans. R. Soc. B* **304**, 575-588 (1984).
3. Feeney, R. E., Burcham, T. S. & Yeh, Y. A. *Rev. Biophys. Chem.* **15**, 59-78 (1986).
4. Duman, J. G. *Cryobiology*, **19**, 613-627 (1982).
5. Raymond, J. A. & DeVries, A. L. *Cryobiology* **9**, 541-547 (1972).
6. Brown, R. A., Yeh, Y., Burcham, T. S. & Feeney, R. E. *Biopolymers* **24**, 1265-1270 (1985).
7. Kerr, W. L., Feeney, R. E., Osuga, D. T. & Reid, D. S. *Cryo-Letters* **6**, 371-378 (1985).
8. Raymond, J. A. & DeVries, A. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2589-2593 (1977).
9. Knight, C. A., DeVries, A. L. & Oolman, L. D. *Nature* **308**, 295-296 (1984).
10. Hallett, J. *J. atmos. Sci.* **21**, 671-682 (1964).
11. DeVries, A. L., Komatsu, S. K. & Feeney, R. E. *J. biol. Chem.* **245**, 2901-2908 (1970).
12. Vlahakis, J. G. & Burduhn, A. J. *A.I.Ch.E. J.* **20**, 581-591 (1974).

Bacterial activity in the warmer, sulphate-bearing, Archaean oceans

Hiroshi Ohmoto & Robert P. Felder

Department of Geosciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

In recent marine sediments, bacterial reduction of seawater sulphate is responsible for the formation of diagenetic sulphides, which are typically strongly depleted in ^{34}S relative to source sulphate, and highly variable in their $\delta^{34}\text{S}$ values. In contrast, the $\delta^{34}\text{S}$ values of Archaean sedimentary sulphides are generally found to be less variable and nearly identical to those of sulphates in the same sedimentary units. This finding has led previous investigators¹⁻⁴ to suggest that either sulphate-reducing bacteria had yet to develop in Archaean time, (especially before ~2.75 billion years ago), and/or Archaean oceans contained much less sulphate ($\ll 1$ mM compared with the present value of 28 mM), implying that the Archaean atmosphere contained much less free oxygen than the present atmosphere. But the sulphur isotope data on Archaean sediments from ~2.6 to 3.5×10^3 Myr old can be better explained if sulphate-reducing bacteria were already active in oceans with temperatures of ~30 to ~50 °C, and containing appreciable amounts (> 1 mM) of sulphate, with $\delta^{34}\text{S}$ values of ~+3%.

The $\delta^{34}\text{S}$ values of ancient seawater sulphate, as indicated by gypsum/anhydrite in ancient evaporites, appears to have fluctuated between ~+10 and +35‰ during the Phanerozoic and Proterozoic eons, with a periodicity of ~20 Myr or longer^{5,6}. The $\delta^{34}\text{S}$ frequency distribution for sulphate minerals from a marine evaporite unit is typically strongly skewed towards isotopically lighter values, with a minimum $\delta^{34}\text{S}$ value that is nearly identical to that of normal seawater SO_4^{2-} (for example, presently +20‰ but ~+10‰ during the Permian period), and with a range of $> 5\%$ ⁵⁻⁷. Such $\delta^{34}\text{S}$ characteristics of evaporite sulphates reflect the very small isotope fractionation factor between gypsum/anhydrite and aqueous SO_4^{2-} , and the very large negative fractionation factor between H_2S generated by SO_4^{2-} -reducing bacteria and SO_4^{2-} . Continuous removal of seawater SO_4^{2-} from a closed evaporitic basin as both gypsum/anhydrite and biogenic sulphides could, therefore, result in a significant increase in the $\delta^{34}\text{S}$ value of the residual seawater SO_4^{2-} .

Sulphate-bearing evaporite bodies are rare in Archaean sediments. Many investigators considered this to be evidence for low- SO_4^{2-} oceans and low-oxygen atmosphere^{1,4}; the occurrence of a few Archaean sulphate bodies was attributed to the local production of SO_4^{2-} from H_2S -bearing ocean water by photosynthetic sulphide-oxidizing bacteria⁸⁻¹⁰. But this theory does not explain why Archaean sulphate minerals with $\delta^{34}\text{S}$ values $< +2\%$ have yet to be found, while sulphide minerals with $\delta^{34}\text{S}$ values $< +2\%$ are so common (Fig. 1). Typical $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values associated with oxidation of H_2S to SO_4^{2-} are within a range of $0 \pm 3\%$ (refs 11, 12), small positive and negative $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values have been found both in the laboratory¹¹ and in nature¹².

Because gypsum and anhydrite are very soluble in low-temperature aqueous solutions (for example, groundwater), the sulphate minerals in older evaporites have less probability of preservation compared to the younger ones¹³. Some sulphate bodies of Archaean age, such as those of the Fig Tree group in Africa and the Warrawoona group in Australia, have been preserved because, during their early diagenetic history, the gypsum was converted to less-soluble barite through interaction with barium-rich hydrothermal fluids^{8,14}. The discovery of ubiquitous occurrences of sulphate-bearing metamorphic minerals (for example, scapolites) in the anhydrite/barite/dolomite-bearing metasediments from the Aldan Shield¹⁵ also suggests that Archaean evaporites were more com-

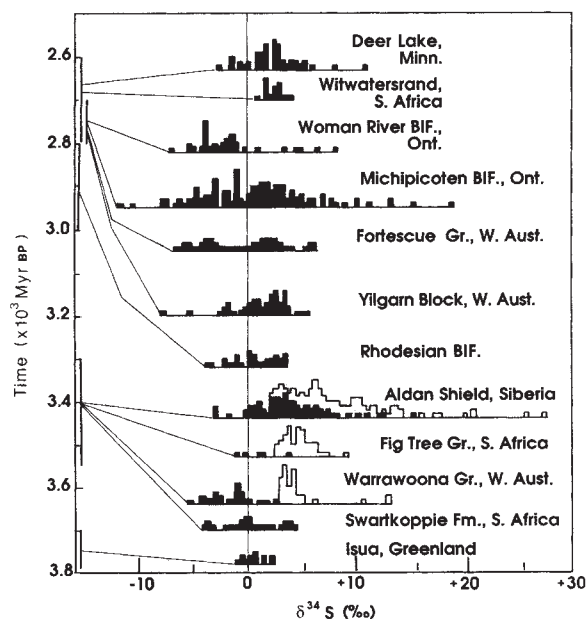


Fig. 1 Comparison of sulphur isotopic composition of Archaean sediments: Isua, Greenland⁴; Swartkoppie Formation, South Africa¹⁰; Warrawoona Group, Western Australia⁸; Fig Tree group, South Africa^{1,8,15}; Aldan Shield, Siberia¹⁵; Rhodesian banded iron formation¹⁹; Yilgarn Block, Western Australia³; Fortescue Group, Western Australia¹⁰; Michipicoten banded iron formation, Ontario, Canada²; Woman River banded iron formation, Ontario, Canada^{2,23}; Witwatersrand, South Africa²⁴; Deer Lake, Minnesota, USA²⁵. The smallest square represents one sample. □, Sulphates; ■, sulphides.

mon than previously believed, and that the oceans were already SO_4^{2-} -rich.

It is significant that the $\delta^{34}\text{S}$ frequency distributions for sulphates from three widely separated areas (Siberia, South Africa and Western Australia) of approximately the same age (~ 3.5×10^3 Myr BP) are very similar; they are strongly skewed toward isotopically lighter values, with minimum $\delta^{34}\text{S}$ values of about +3‰ and ranges of $> 5\%$. A comparison of these data with those of younger evaporites (discussed above) suggests that, ~ 3.5×10^3 Myr ago, the oceans already contained appreciable amounts of SO_4^{2-} with $\delta^{34}\text{S}$ values of about +3‰, and sulphate-reducing bacteria were already active to produce significant sulphur isotope variation of SO_4^{2-} in closed basins. Sufficient organic matter (derived from dead biota) readily metabolized by sulphate-reducing bacteria would have been available; the abundance of CO_2 -metabolizing organisms has been indicated by carbon isotope studies of Archaean sediments¹⁰. The exact SO_4^{2-} contents in these oceans are difficult to estimate, but they were probably more than ~1 mM ($> 1/30$ of the present ocean value) to account for large sulphate bodies, such as the barite-bearing unit in the Fig Tree group which is up to 20 m thick with a minimum areal extent of 500 m $\times$ 10 km (refs 1, 14). We now consider the validity of the above suggestions by examining the sulphur isotope data on diagenetic sulphide minerals.

Most sulphide minerals in recent marine sediments form during shallow burial, within ~200 cm of the seafloor^{6,16-18}, and their $\delta^{34}\text{S}$ values are controlled largely by three parameters: (1) the $\delta^{34}\text{S}$ value of local seawater SO_4^{2-} , which may be influenced by the open- or closed-system nature of the basin; (2) the kinetic isotope fractionation factor accompanying bacterial sulphate reduction (the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ value); and (3) the open or closed nature of the sediment system (not of the overlying seawater system) with respect to SO_4^{2-} , which is determined by the relative rates of SO_4^{2-} -supply and SO_4^{2-} -reduction (Fig. 2). The uppermost

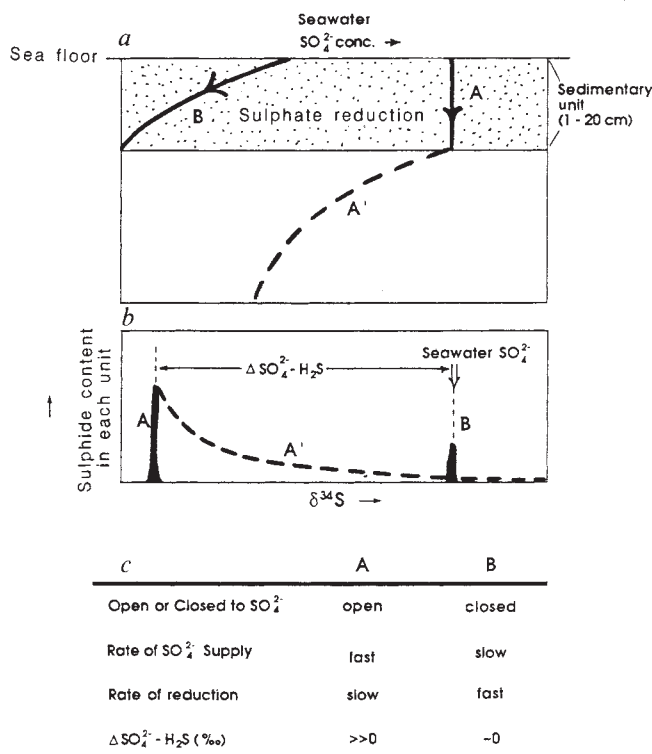


Fig. 2 *a*, A schematic representation of the change in the SO_4^{2-} contents of pore fluids in three different systems: A (open to SO_4^{2-}), B (closed to SO_4^{2-}) and A' (partially closed to SO_4^{2-}). *b*, A schematic comparison of the sulphide contents and the $\delta^{34}\text{S}$ values of sulphide minerals formed in the above three systems. *c*, A table summarizing the relationships among the open or closed nature of a sedimentary system, rate of SO_4^{2-} supply, rate of SO_4^{2-} reduction, and $\delta^{34}\text{S}$ characteristics of sulphides.

~20 cm of recent sediments, where about one half of the sedimentary sulphides are formed^{6,16-18}, can be considered a system open to SO_4^{2-} (SO_4^{2-} -supply rate $\gg \text{SO}_4^{2-}$ -reduction rate). The SO_4^{2-} content of the pore fluids remains essentially the same as that of overlying seawater, despite intensive bacterial sulphate reduction (curve A, Fig. 2), because SO_4^{2-} is supplied at fast rates due to the churning of sediments by burrowing organisms. The sulphides formed in such an open system possess nearly uniform $\delta^{34}\text{S}$ values, which differ from those of seawater SO_4^{2-} (presently $+20\%$) by the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ value. The $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values average about 50‰ in recent sediments, but vary between ~70 and ~20‰, depending on sedimentary environments^{6,17,18}.

At depths between ~20 and ~200 cm in recent marine sediments, where the remaining one half of the sedimentary sulphides are fixed, SO_4^{2-} is replenished by diffusion, but the SO_4^{2-} -supply rate is less than SO_4^{2-} -reduction rate (the system is semi-closed to SO_4^{2-})^{6,16-18}. In such systems, the H_2S and SO_4^{2-} in the pore fluids continuously decrease in concentration but increase in $\delta^{34}\text{S}$ value with depth (curve A', Fig. 2). Therefore, during the continuous burial of a packet of sediments from the seafloor to depths of ~200 cm, the environment for the packet of sediments changes from an open system to a semi-closed system with respect to SO_4^{2-} : H_2S with increasing $\delta^{34}\text{S}$ values reacts continuously with iron in sediments and is added to the sulphide minerals formed earlier in the open system. These are the reasons for the following important characteristics^{6,17,18} of sulphides in recent marine sediments: (1) the isotope variability is typically around 20‰ within a sedimentary sequence; (2) the $\delta^{34}\text{S}$ distribution curve is strongly skewed toward isotopically lighter values (a combination of curves A and A' in the middle

figure of Fig. 2); and (3) the mean $\delta^{34}\text{S}$ value for samples from an area lies typically within a range of $-25 \pm 20\%$ (45 $\pm 20\%$ less than the $\delta^{34}\text{S}$ of seawater SO_4^{2-}). These three characteristics have also been recognized in many marine sedimentary units of Phanerozoic and Proterozoic age⁶, although the $\delta^{34}\text{S}$ values of seawater SO_4^{2-} varied between $\sim +10$ and $+35\%$.

In contrast, sulphides of Archaean age (Fig. 1) have the following sulphur isotope characteristics: (1) the variability may possibly be smaller: the $\delta^{34}\text{S}$ variation has been suggested to be less than 10‰ in a sampled area^{1,4,8,10,19}, however, this suggestion may be misleading. The variation in the $\delta^{34}\text{S}$ values in Fig. 1 appears to be proportional to the number of samples analysed; where a large number of samples were analysed (for example, Aldan Shield and Michipicoten Iron Formation), the $\delta^{34}\text{S}$ variation is comparable to that of recent marine sediments; (2) the $\delta^{34}\text{S}$ frequency distribution is not strongly skewed toward lighter values (that is, either normal distributions or skewed toward heavier or slightly lighter values); and (3) the mean $\delta^{34}\text{S}$ value for samples from each region lies within a narrow range of $+1 \pm 2\%$, which is only ~2‰ less than the $\delta^{34}\text{S}$ value of Archaean seawater SO_4^{2-} proposed above. These characteristics of Archaean sulphides can be explained by the activity of sulphate-reducing bacteria, if: (1) the $\delta^{34}\text{S}$ value of seawater SO_4^{2-} remained relatively constant at about $+3\%$ for the period between ~3.5 and ~2.6 $\times 10^3$ Myr BP; (2) the kinetic isotope effect accompanying bacterial sulphate reduction (the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values) was within a range of $10 \pm 5\%$; and (3) the sedimentary environments were more like a system closed to SO_4^{2-} , rather than open to SO_4^{2-} . When the system is completely closed to SO_4^{2-} (SO_4^{2-} -supply rate $\ll \text{SO}_4^{2-}$ -reduction rate), SO_4^{2-} is quantitatively converted to sulphide within the system, thus, regardless of the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ value, the $\delta^{34}\text{S}$ value for sulphides becomes identical to the initial $\delta^{34}\text{S}$ value of SO_4^{2-} (curve B, Fig. 2). In other words, with decreasing openness of the system with respect to SO_4^{2-} , the $\delta^{34}\text{S}$ frequency distribution for sulphides continuously changes from strongly skewed towards less positive values to strongly skewed towards more positive values, while the mean $\delta^{34}\text{S}$ value approaches that of seawater SO_4^{2-} .

From the data on chemical and isotope characteristics of recent marine sediments, it has been suggested^{6,17} that the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values are inversely proportional to the logarithm of the rates of SO_4^{2-} reduction; the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values may decrease from ~60‰ (typically in deep-sea sediments) to ~30‰ (typically in near-shore sediments) with an increase in the reduction rates by about two log units. A close examination of experimental data on bacterial sulphate reduction (Fig. 3) also indicates that the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values are inversely proportional to the logarithm of the rate of sulphate reduction.

In Archaean marine sediments, the rate of SO_4^{2-} -supply must have been much less than that in the modern system, because of the absence of burrowing organisms. Therefore, in contrast to recent sediments, the general tendency for Archaean sediments would have been to behave more like a closed system with respect to SO_4^{2-} ; this tendency would have been enhanced if the SO_4^{2-} -reduction rates in Archaean sediments were much higher than those in recent sediments.

The rate of sulphate reduction (and the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values) depends on various parameters, such as bacterial species, types and concentrations of electron donors (that is, nutrients for bacteria) and temperature^{11,20-22}. In Archaean time, *Desulfovibrio desulfuricans* may or may not have been the most important sulphate-reducing bacteria; their metabolism may also have been different from the recent species. But if the metabolic response of Archaean sulphate-reducing bacteria to a change in temperature was similar to that of *Desulfovibrio desulfuricans* used in the experiments of Fig. 3, the only apparent mechanism to significantly increase the SO_4^{2-} -reduction rates and decrease the $\Delta_{\text{SO}_4-\text{H}_2\text{S}}$ values to $10 \pm 5\%$ worldwide in the Archaean would be an increase in the ocean temperatures to ~30–~50 °C (compared to the present 4 °C). The maximum temperature of experiments

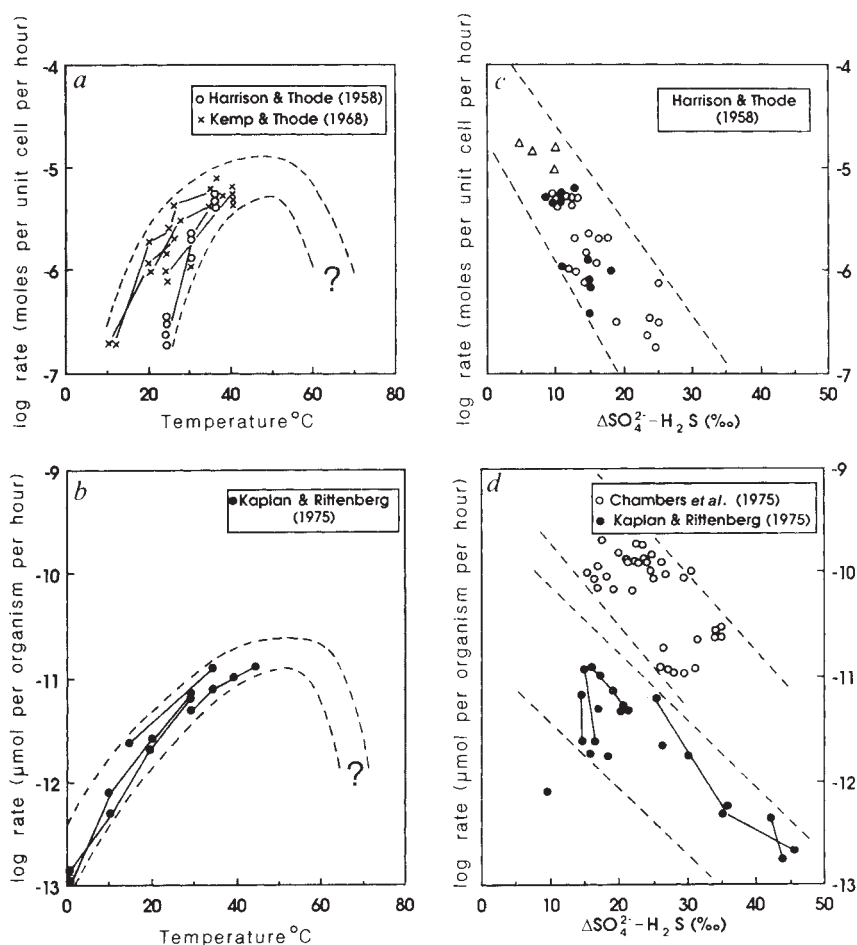


Fig. 3 *a* and *b*, The rate of bacterial sulphate reduction and the kinetic isotopic effects observed in various laboratory experiments^{11,20,22} are compared using *Desulfovibrio desulfuricans*, lactate and other organic compounds as electron donors, and solutions with SO_4^{2-} contents of 10 to 100 mM. The various groups of researchers used different strains of bacteria, and normalized their results to different bacterial units (for example, rates per organism or per colony of bacteria); their rate values are not directly comparable, but the magnitudes of change in the rates are. *c* and *d*, The relationships between the rates of bacterial reduction and temperature observed in some laboratory experiments^{11,20,21} are shown. Solid lines in the figures connect data points obtained in experiments carried out under similar conditions.

in Fig. 3 was 45 °C, where the rates are about two orders of magnitude greater than those at ~4 °C. Although *Desulfovibrio desulfuricans* itself does not grow above ~40 °C (ref. 26), there are modern sulphur-reducing bacteria that grow best at 90 °C (ref. 27), and there is no *a priori* reason why Archaean sulphate-reducers could not have lived at high temperatures.

The above reasoning further leads to the suggestion that a continuous decrease in the ocean temperatures from Archaean to Phanerozoic was the principal reason why the $\delta^{34}\text{S}$ of both the sulphides and sulphates in the younger marine sediments shifted further away from ~+3‰ and became more variable. The development of burrowing fauna, which also tends to make sedimentary systems more open to SO_4^{2-} , would have further enhanced the above $\delta^{34}\text{S}$ trends in Phanerozoic period.

We do not imply that sulphide minerals in Archaean sediments were all biogenic, nor that bacterial activity is the sole mechanism for sulphur isotope variation in nature. Some Archaean sulphides and sulphates, especially those in veins and massive sulphide deposits, were formed probably by hydrothermal processes; their $\delta^{34}\text{S}$ variation may have been caused by a variation in the $\text{SO}_4^{2-}/\text{H}_2\text{S}$ ratio of hydrothermal solutions^{3,12,19}. Some investigators^{3,19} have attributed the variation in the $\text{SO}_4^{2-}/\text{H}_2\text{S}$ of Archaean hydrothermal solutions to that in the oxidation of H_2S -bearing magmatic fluids. We will demonstrate (H.O., in preparation) that the Archaean hydrothermal solutions with variable $\text{SO}_4^{2-}/\text{H}_2\text{S}$ ratios were probably of seawater origin: some H_2S was produced by non-bacterial reduction of sulphate-bearing seawater during its deep circulation through warmer crustal rocks. Indeed, the sulphur isotope variation in some Archaean hydrothermal deposits can be used as another line of evidence for sulphate-bearing Archaean oceans.

This work was supported by the NSF. We thank H. D. Holland

for discussion, C. J. Kaiser, L. R. Kump, I. B. Lambert, P. A. Trudinger, T. H. Donnelly and M. L. Coleman for helpful comments on an earlier manuscript.

Received 15 December 1986; accepted 29 June 1987.

- Perry, E. C., Monster, J. & Reimer, T. O. *Science* **171**, 1015-1016 (1971).
- Goodwin, A. M., Monster, J. & Thode, H. G. *Econ. Geol.* **71**, 870-891 (1976).
- Donnelly, T. H. *et al. J. geol. Soc. Aust.* **24**, 409-420 (1977).
- Monster, J. *et al. Geochim. cosmochim. Acta* **43**, 405-413 (1979).
- Claypool, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. *Chem. Geol.* **28**, 199-260 (1980).
- Ohmoto, H., Kaiser, C. J. & Geer, K. A. *Spec. Publ. geol. Soc. Aust.* (in the press).
- Thode, H. G. & Monster, J. in *Fluids in Subsurface Environments* (eds Young, A. & Galley, J. E.) 367-377 (Am. Ass. Petrol. Geol. Memoir 4, 1965).
- Lambert, I. B., Donnelly, T. H., Dunlop, J. S. R. & Groves, D. I. *Nature* **276**, 808-811 (1978).
- Broda, E. *The Evolution of Bioenergetic Processes* (Pergamon, Oxford, 1975).
- Schidlowski, M., Hayes, J. M. & Kaplan, I. R. in *Earth's Earliest Biosphere, Its Origin and Evolution* (ed. Schopf, J. W.) 149-186 (Princeton University Press, 1983).
- Kaplan, I. R. & Rittenberg, S. C. *J. gen. Microbiol.* **34**, 195-212 (1964).
- Ohmoto, H. & Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 509-576 (Wiley, New York, 1979).
- Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton University Press, 1984).
- Heinrichs, T. K. & Reimer, T. O. *Econ. Geol.* **72**, 1426-1441 (1977).
- Vinogradov, V. I., Reimer, T. O., Leites, A. M. & Smelov, S. B. *Lithol. Miner. Resources* **11**, 407-420 (1977).
- Berner, R. A. *Geochim. cosmochim. Acta* **48**, 605-617 (1984).
- Goldhaber, M. B. & Kaplan, I. R. *Soil Sci.* **119**, 42-55 (1975).
- Goldhaber, M. B. & Kaplan, I. R. in *The Sea Vol. 5, Marine Chemistry* (ed. Goldberg, E. D.) 569-655 (Wiley, New York, 1974).
- Fripp, R. E. P., Donnelly, T. H. & Lambert, I. B. *Spec. Publ. geol. Soc. S. Afr.* **5**, 205-208 (1979).
- Harrison, A. G. & Thode, H. G. *Trans. Faraday Soc.* **54**, 84-92 (1958).
- Kemp, A. L. W. & Thode, H. G. *Geochim. cosmochim. Acta* **32**, 71-91 (1968).
- Chambers, L. A., Trudinger, P. A., Smith, J. W. & Burns, M. S. *Can. J. Microbiol.* **21**, 1602-1607 (1975).
- Thode, H. G. & Goodwin, A. M. *Precamb. Res.* **20**, 337-356 (1983).
- Hoefs, J., Nielsen, H. & Schidlowski, M. *Econ. Geol.* **63**, 975-977 (1968).
- Ripley, E. M. & Nicol, D. L. *Geochim. cosmochim. Acta* **45**, 839-846 (1981).
- Postgate, J. R. *The Sulphate-reducing Bacteria* 2nd edn (Cambridge University Press, 1984).
- Segerer, A., Stetter, K. O. & Klink, F. *Nature* **313**, 787-789 (1985).

Altered geometry of webs in spiders with regenerated legs

Fritz Vollrath

Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

The form of a spider's web can be seen as a representation of inherited rules of behaviour, although the geometry of the orb web is not rigidly predetermined. Examining the web as a product of orientation processes is a novel and powerful tool for analysing the measuring functions and decision rules of the spider's web-building behaviour. The garden spider *Araneus diadematus* Clerk orients the main (capture) spiral of the web according to measurements taken by particular legs. This is apparent when a measuring leg has been regenerated, because the reduced length of the regenerate is reflected in the geometry of the capture spiral. The orientation of the auxiliary spiral (used for support during construction), on the other hand, is not affected and I conclude that the two spirals, both of which are present in each web, are constructed according to different behavioural rules.

The spider *Araneus diadematus* builds its orb web in three distinct phases. The spider first joins frame threads and radii, then links the radii with a coarse and nonsticky auxiliary spiral (from the centre outwards), and finally puts down the fine and sticky capture spiral (from the periphery inwards). The auxiliary spiral is removed during construction of the capture spiral, suggesting that it functions as a temporary stabilizing platform^{1,2}. The differences in function (framework or trap) between the two spirals are reflected in their material (resilient silk and elastic silk) as well as in their geometry (non-arithmetic spirals compared to arithmetic ones). Whereas the spacing of the auxiliary spiral increases with each joint, the capture spiral is evenly spaced throughout most of the web². This observation suggests that two different rules of orientation might determine the placement of each spiral because in an arithmetic spiral the distances between turns are constant, whereas in a non-arithmetic spiral the angles between spiral and radii are constant³.

This hypothesis was tested by analysing the geometry of webs built by spiders which have regenerated certain legs. All webs were built under standard laboratory conditions and the webs were photographed, digitized and analysed by computer^{2,4}. *Araneus* normally regenerates a lost leg during the following moult and uses the new appendages soon afterwards (manuscript in preparation), so that the spider can build a nearly perfect web only 24 h after all legs on one side of the body have been regenerated. Because injured or trapped legs are naturally shed (appendotomized) at the coxa-trochanter joint⁵, spiders with all combinations of normal and short legs can easily be obtained. Such 'unbalanced' individuals provide valuable infor-

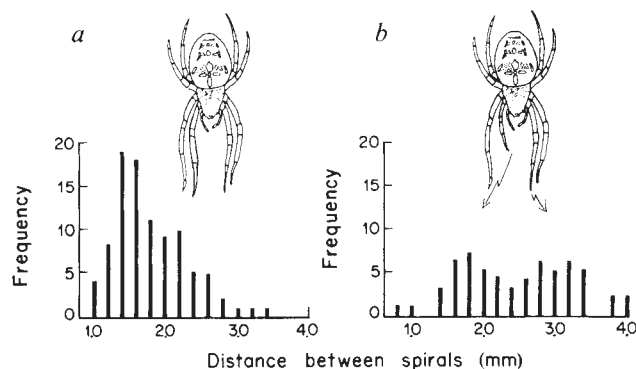


Fig. 1 Capture spiral: histograms of the distances between turns of the capture spiral (mesh size) in the lower part (South) of *Araneus diadematus* webs. Representative for 20 measurements of normal webs: *a*, distribution in a web built by a spider (A035) with normal legs; *b*, distribution in a web built by a spider (A157) which had regenerated the first right leg and used either the regenerated or the normal leg for measuring. Although *b* is also a single representative example, the statistics on the effect of leg length on capture spiral spacing are based on the analysis of 10 webs of 10 spiders, each web providing ~60 measurements. For this analysis the two peaks of the bimodal distributions for each web were determined by eye and were analysed as pairs in a Wilcoxon test ($N=10$ pairs), giving a highly significant result ($P<0.001$). The means and standard deviations for spiral spacing (x) were: regenerated legs, $\bar{x}=18.7\pm3.7$ mm; normal legs, $\bar{x}=28.8\pm5.2$ mm.

mation about the measuring and executing function (input and output) of each leg. Spiders that had regenerated one, several, or all legs on one side of the body were used in preference because they combined control and experimental conditions (normal and short legs) in the same animal.

It has been suggested^{6,7} that the legs on one side of the body (ipsilateral) measure the distance to the previous turn of the spiral (the auxiliary spiral proceeds outwards, the capture spiral inwards). This hypothesis can be tested conclusively using spiders with regenerated legs because in this case the length of the measuring leg should be reflected in the spacing of the spiral. During construction of the capture spiral, the spider may often turn around in its track. Then the control and the experimental legs face outwards (from the centre of the web) alternately, and thus the measuring and executing (manipulative) function of all eight legs can be displayed in a single capture spiral (Fig. 1*b*). The animal does not 'loop' while building auxiliary spirals, therefore auxiliary spiral geometry was studied in different webs of the same individuals (Fig. 2).

The results confirm that the capture spiral is fixed to the radius by legs III and IV, and that the point of placement is found by leg I (or II, if I is missing). When all legs facing the previous

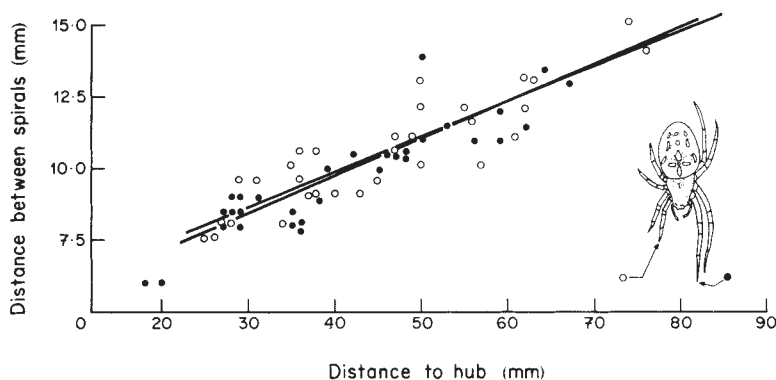


Fig. 2 Auxiliary spiral: the regression for spiral spacing of the auxiliary spirals in two webs built by the same spider (A10). All four legs on the right body side were regenerated. The different symbols represent the measurements for the two spirals. Either the regenerated legs were ipsilateral with the previous spiral turn (white dots) or the normal legs were ipsilateral (black dots). For the statistics 18 webs of 2 spiders (A10 and A9) were analysed, each spider providing 6 webs (3 with short legs and 3 with long legs ipsilateral). Each web provided ~30 points for a regression analysis, all correlation coefficients being highly significant. In spider A10 all six regressions overlapped completely, tested for the equality of slopes¹³. In spider A9 the slopes of the six regressions had some scatter, the slope being a measure for the mesh size of a non-arithmetic spiral. Detailed analysis did not show any significant differences between the spiral slopes due to treatment (long versus short legs), tested with the GTZ multiple comparison test.¹³

spiral turn are shorter, a significant change in the spacing of the capture spiral occurs. However, the same result is also obtained when only the first (I) of the legs on this side is shorter (Fig. 1b). It follows that the first leg on the side nearest the previous turn measures the distance to the previous spiral-turn, and that it thus determines the spacing of the capture spiral. Spacing of the auxiliary spirals showed no effect of leg length, not even when all four legs on this side were shorter (Fig. 2). This suggests that the auxiliary spiral is not spaced according to distance measurements taken by any leg (I to IV). It is possible, but not probable⁵, that the distance is measured by the whole body. A more likely alternative hypothesis² is that the auxiliary spiral is oriented according to measurements of the angle formed between each spiral section and the adjoining radius; these angular measurements would not be affected by the length of the measuring legs.

I conclude that the behavioural rules which determine the pattern of each spiral are different. This would explain the results of earlier experiments on the importance of gravity as a compass reference for the spider's orientation during web construction^{4,7}. These experiments, using a clinostat, showed that rotation in the vertical plane can temporarily disorient the web-building spider, resulting in tangled webs with dramatically deranged capture spirals but with normal auxiliary spirals⁷.

Araneus and some related orb weavers stand out among the spiders in their ability to use immediately newly regenerated legs (manuscript in preparation). Most spiders studied so far use regenerated legs only after additional moults, which is presumably the ancestral condition. One group of orb weavers (*Nephila*, *Zygiella*) have evolved a biochemical mechanism which actively suppresses regeneration localized at the coxal joint of appendotomy; these species cannot use a leg that is experimentally induced to regenerate distal to the coxa. This suggests that the ability to coordinate the placement of regenerates with normal legs may require special adaptations of the 'program' which controls orb-web-building behaviour; such 'subroutines' seem to have evolved in *Araneus* but not in *Nephila*, which instead foregoes the benefits of regeneration.

An orb web is the outcome of orientation behaviour. The way in which the spider uses a regenerated leg enables us to infer some of the orientation parameters which determine the animal's position and path. This deductive approach allows us to test the theory that some spiders use kinaesthetic or idiothetic cues to navigate in their webs^{11,12}, that is that *en route* proprioceptive information is either used to give the spider a series of positions¹² or is integrated into a mental map^{13,14}. This theory of arthropod navigation by mental maps is controversial^{15,16}; the method described might be extended to other arthropods.

I thank Dr Richard Dawkins for support throughout the study, which was financed by the SERC. Alan Solomon helped with the statistics. Drs M. Dawkins, R. Dawkins, C. Graham and P. Harvey kindly read the manuscript and suggested improvements.

Received 12 May; accepted 22 May 1987.

1. Eberhard, W. G. *Evolution* **36**, 1067-1095 (1982).
2. Vollrath, F. & Mohren, W. *Naturwissenschaften* **72**, 666-667 (1985).
3. D'Arcy Thompson, W. *On Growth and Form* Vol. 2, (Cambridge University Press, 1942).
4. Vollrath, F. *J. Comp. Physiol. A* **152**, 275-280 (1986).
5. Roth, V. D. & Roth, B. M. *Bull. Br. arachnol. Soc.* **6**, 137-146 (1984).
6. Peters, H. M. Z. *Morph. Ökol. Tiere* **32**, 613-649, 33, 128-150 (1937).
7. Vollrath, F. *J. comp. Physiol. A* (in the press).
8. Ruhland, M. thesis, Univ. Konstanz (1976).
9. Bonnet, P. thesis, Univ. Toulouse (1930).
10. Randall, J. B. *Wilhelm Roux Arch. dev. Biol.* **190**, 230-232 (1981).
11. Goerner, P. & Claas, B. in *Neurobiology of Arachnids* (ed. Barth, F. G.) 275-297 (Springer, Berlin, 1985).
12. Mittelstaedt, H. in *Neurobiology of Arachnids* (ed. Barth, F. G.) 298-316 (Springer, Berlin, 1985).
13. Hill, D. E. *Behav. Ecol. Sociobiol.* **5**, 301-322 (1979).
14. Gould, J. L. in *The Biology of Learning* (eds Marler, P. & Terrace, H.) 149-180 (Springer, Berlin, 1984).
15. Wehner, R. in *Neuroethology and Behavioural Physiology* (eds Huber, F. & Markl, H.) 366-381 (Springer, Berlin, 1983).
16. Gould, J. L. *Science* **232**, 861-863 (1986).

Genomic imprinting determines methylation of parental alleles in transgenic mice

Wolf Reik, Andrew Collick, Michael L. Norris, Sheila C. Barton & M. Azim Surani

Department of Molecular Embryology, Institute of Animal Physiology and Genetics Research, Babraham, Cambridge CB2 4AT, UK

Mouse embryogenesis relies on the presence of both the maternal and the paternal genome for development to term^{1,2}. It has been proposed that specific modifications are imprinted onto the chromosomes during gametogenesis³; these modifications are stably propagated⁴, and their expression results in distinct and complementary contributions of the two parental genomes to the development of the embryo and the extraembryonic membranes^{5,6}. Genetic data further suggest that a substantial proportion of the genome could be subject to chromosomal imprinting⁷, the molecular nature of which is unknown. We used random DNA insertions in transgenic mice to probe the genome for modified regions. The DNA methylation patterns of transgenic alleles were compared after transmission from mother or father in seven mouse strains carrying autosomal insertions of the same transgenic marker. One of these loci showed a clear difference in DNA methylation specific for its parental origin, with the paternally inherited copy being relatively undermethylated. This difference was observed in embryos on day 10 of gestation, but not in their extraembryonic membranes. Moreover, the methylation pattern was faithfully reversed upon each germline transmission to the opposite sex. Our findings provide evidence for heritable molecular differences between maternally and paternally derived alleles on mouse chromosomes.

Heterozygous transgenic mice with insertions of a chloramphenicol acetyl transferase gene (CAT) linked to an immunoglobulin heavy-chain enhancer (IgH-E; ref. 8) were bred with wild-type animals of the same genetic background (C57BL/6J × CBA/Ca)F₁ (Fig. 1). Fetuses were dissected on day 10 of gestation because functional differences between parental chromosomes are most prominent at this time^{5,6} and DNA was isolated from the embryo and from the extraembryonic membranes (the trophoblast and yolk sac). The effect of the parental origin of the transgene on its state of methylation was assessed by comparing methylation patterns after maternal and paternal transmission in each strain, using the methylation-sensitive enzyme *Hpa*II (Fig. 1). In six strains, the transgene was found to be highly methylated in the embryo and undermethylated in the trophoblast and yolk sac, irrespective of parental origin (data not shown). This pattern of lineage-specific methylation has also been observed with endogenous DNA sequences, both repetitive and unique^{9,10}. Hence, the CAT-IgH-E construct does not appear to contain specific sequences that act as a signal for differential modification in male and female germ cells. However, in embryos of one of the transgenic strains, designated CAT17 (ref. 8), we found a striking difference between male- and female-derived alleles; a *Bam*HI fragment of 4.3 kilobases (kb) was virtually unmethylated when derived from the father, but highly methylated when derived from the mother (Fig. 2a). No difference was seen in the extraembryonic tissues. Digestion with *Hpa*II alone released a fragment of 3.9 kb diagnostic for paternal transmission of the DNA that was absent after maternal transmission, confirming demethylation of the DNA derived from the father (Fig. 2b).

One of the most important requirements for the mechanism of chromosomal imprinting is the reversibility of the modification in germ cells of the opposite sex³. The reversibility of the methylation pattern in strain CAT17 was tested as the

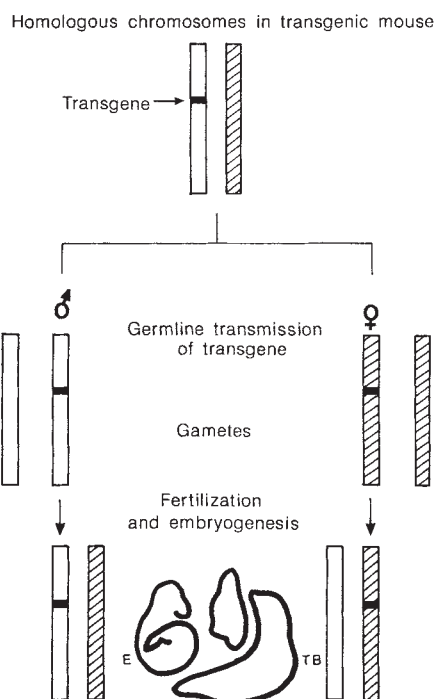


Fig. 1 Maternal and paternal transmission of transgenic markers into embryos. Mice from seven different strains transgenic for CAT-IgH-E, with copy numbers ranging from 1 to 6 mostly in head-to-tail arrangements⁸, were bred with wild type animals of the same genetic background (C57BL/6J × CBA/Ca)F₁. Heterozygous males and females derived from these crosses were bred with wild type animals. For each strain, DNA from heterozygous embryos (E) on day 10 of gestation with maternal or with paternal inheritance of the transgene and from corresponding extraembryonic tissues (TB) was restricted with *Eco*RI (two sites in CAT-IgH-E⁸) and/or *Bam*HI (1 site in CAT-IgH-E⁸) and in double digests with *Eco*RI-*Hpa*II and *Bam*HI-*Hpa*II. *Eco*RI-*Msp*I and *Bam*HI-*Msp*I digests were included as a control for complete undermethylation of five *Hpa*II sites present in CAT-IgH-E. For most strains more than one embryo was analysed for each mode of transmission (maternal, paternal) to exclude variability between embryos. In some cases embryos were pooled in pairs before DNA extraction. Open bars, paternal chromosomes; hatched bars, maternal chromosomes.

Methods. Pregnant females were autopsied on day 10 of gestation (day 1 = day of plug detection) and embryos, yolk sacs and trophoblasts were isolated. The yolk sac and trophoblast were pooled; yolk sac mesoderm was not analysed separately, although methylation levels may be expected to be high, as in the embryo. But the DNA from yolk sac mesoderm represents only a minor fraction of the total DNA from extraembryonic tissues and was therefore included. Tissues were homogenized in phosphate-buffered saline (PBS), cells were resuspended in 75 mM NaCl, 25 mM EDTA pH 8, an equal volume of 10 mM Tris pH 8, 10 mM EDTA, 1% SDS, 200 µg ml⁻¹ Proteinase K was added, and the lysates were incubated at 50 °C for 3 h. Digests were extracted with phenol and chloroform, isopropanol-precipitated and the DNA was redissolved in TE. Restriction enzymes were used under conditions recommended by the suppliers. Samples of restriction digests (1 µg) were electrophoresed through 1% agarose gels, alkali-transferred to Gene Screen plus membranes and hybridized with CAT-specific and IgH-E-specific fragments labelled by oligonucleotide priming²². Final stringency of washing was 0.1 × SSC, 1% SDS, 68 °C. Completion of digests was controlled by inclusion of phage λ DNA, crucial digests were controlled by rehybridization of the filter to a mitochondrial probe, p501-1, which results in identical *Hpa*II and *Msp*I patterns in all tissues²³.

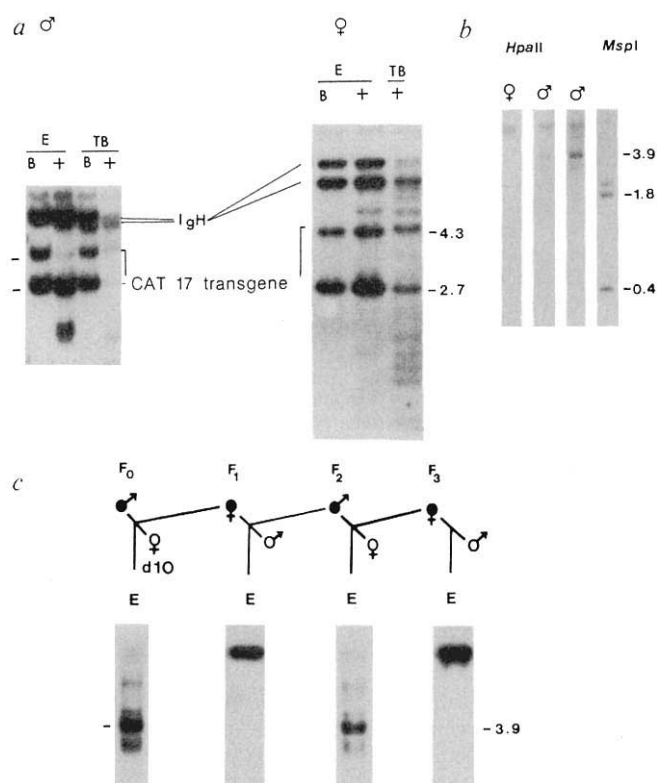


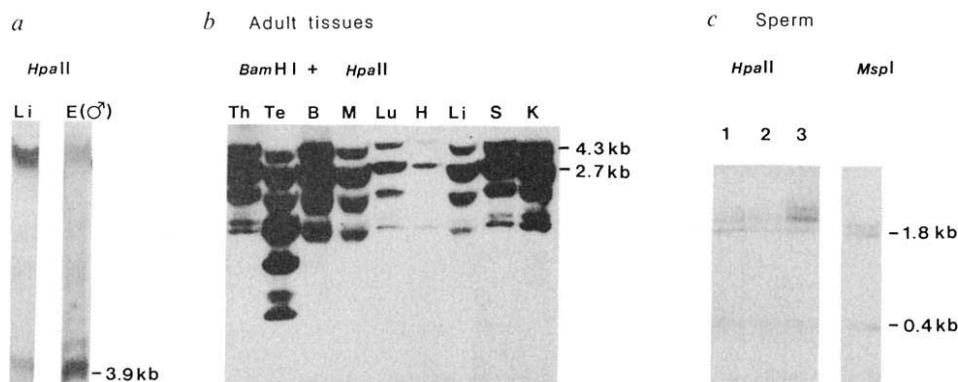
Fig. 2 Methylation of maternal and paternal alleles in strain CAT 17. Embryos and extraembryonic membranes were analysed individually on day 10 of gestation as described in legend to Fig. 1. *a*, DNA from embryo (E) and yolk sac and trophoblast (TB) were restricted with *Bam*HI (B) and with *Bam*HI-*Hpa*II (+), and hybridized with a 0.9-kilobase(kb) *Xba*I fragment that contains the IgH enhancer⁸. Paternal (♂) and maternal (♀) transmission are compared. The transgene consists of two *Bam*HI fragments, that of 2.7 kb resulting from head-to-tail arrangement of 3 copies of CAT-IgH-E and that of 4.3 kb presumably extending into flanking mouse sequences (not shown). Two polymorphic *Bam*HI variants of the heavy-chain locus (IgH) come from the C57BL/6J strain and from the CBA/Ca strain (W.R., unpublished) and segregate randomly in our transgenic strains. In the analysis of paternal transmission (♂), the embryo (E) and the extraembryonic tissues (TB) were from different conceptuses. Due to the limited resolution of the gel to the left (♂) the methylation patterns of extraembryonic tissues are difficult to compare (♂TB+ with ♀TB+). Analysis of a number of fetuses with different digests confirmed that the patterns were identical (not shown). *b*, DNA from 3 embryos, after maternal (♀) and paternal (♂) transmission of the transgene, was restricted with *Hpa*II and hybridized with a 1.8-kb *Hind*III-*Bam*HI fragment specific for CAT⁸. An *Msp*I digest was included as a control for complete undermethylation of CAT 17. *c*, Alternate transmission of the transgene through male and female germlines. Animals heterozygous for CAT 17 (♂, ♀), starting with the male founder, were bred with wild type animals (♂, ♀) to derive heterozygous animals of the next generation, as well as fetuses on day 10 (d10) for methylation analysis in embryo (E) and extraembryonic membranes. Restriction of embryonic DNA was with *Hpa*II, hybridization was with a CAT-specific fragment; DNA from extraembryonic tissues not shown. Fragment sizes are indicated in kilobases. The heterozygous animals from this pedigree were also monitored for expression of the transgene by enzymatic assay for CAT activity in spleen and testis (one F₂ male). CAT activity was not detectable (not shown). Similar results were consistently obtained with about 90 embryos from 20 heterozygous parents analysed. See Fig. 1 legend for methods.

transgene was transmitted from the founder animal to the F₃ generation alternately through the male and the female germline (Fig. 2c). Animals from this pedigree were bred with wild type animals to derive day 10 embryos from each generation. Methylation of the transgenic DNA switched from the pattern diagnos-

tic for paternal transmission to the maternal transmission pattern and back in each subsequent generation (Fig. 2c). Hence, the methylation status specific for the parental origin of the DNA was clearly reversible upon each germline passage to the opposite sex, and was heritable for at least three generations.

Fig. 3 Levels of methylation in adult tissues and in sperm of animals from strain CAT 17. A heterozygous male that inherited the CAT 17 locus from his mother was killed at the age of 10 wk and DNA was collected from different organs as described in the legend to Fig. 1. *a*, DNA from liver from this animal (Li) and from a day 10 embryo that inherited the transgene from the father (E(δ)) was digested with *Hpa*II, electrophoresed on a 0.5% agarose gel and hybridized with a CAT-specific probe. The 3.9-kb *Hpa*II fragment characteristic for male transmission is indicated. The difference in DNA methylation seen between adult liver after maternal transmission and the embryo after paternal transmission was also observed with adult tissues, for example tail, after both modes of transmission (not shown). *b*, DNA from thymus, Th; testis, Te; brain, B; striated muscle, M; lung, Lu; heart, H; liver, Li; spleen, S and kidney, K; was restricted with *Bam*HI and *Hpa*II, electrophoresed through a 0.8% agarose gel and hybridized with a CAT-specific probe. The two *Bam*HI fragments of the CAT 17 locus of 2.7 kb and 4.3 kb are indicated. *c*, DNA from sperm from three individual adult heterozygous males of strain CAT 17 (1, 2, 3). Animals no. 1 and no. 3 inherited their transgene from the father, whereas animal no. 2 inherited it from the mother. The sperm DNA was digested with *Hpa*II, electrophoresed through a 0.8% agarose gel and hybridized with a CAT-specific probe. DNA from CAT 17 digested with *Msp*I was included as a control for complete undermethylation. Fragment sizes are indicated.

Methods. Sperm was collected from the epididymis and the ductus deferens and DNA was isolated as described in Fig. 1 legend, except that 70 mM β -mercaptoethanol and 0.5 mg ml⁻¹ Proteinase K was used in the lysis solution. All other methods as described in legend to Fig. 1.



Maternally and paternally derived chromosomes when segregated in parthenogenetic-androgenetic chimaeras apparently confer different spatial specificities on cells, such that on day 10 of gestation parthenogenetic cells are mainly confined to the embryo and the yolk sac whereas androgenetic cells are found in the trophoblast and yolk sac⁶. We note in this respect that the effect of methylation imprinting in strain CAT 17 is confined to the embryos and is not detected in the extraembryonic tissues. To establish whether the difference between parental alleles is segregated into a specific lineage at developmental stages later than day 10 we analysed DNA from different organs from an adult heterozygous male that inherited his CAT-IgH-E copy from his mother (Fig. 3). Methylation levels in all organs tested, with the exception of testis, were identical and, although slightly lower than in day 10 embryos with maternal transmission of the transgenic allele, they were clearly higher than in embryos with paternal transmission (Fig. 3*a, b*). DNA from mature testis, which consists primarily of germ cells, was strikingly undermethylated (Fig. 3*b*). Indeed, the CAT 17 locus was virtually unmethylated in mature spermatozoa (Fig. 3*c*). Undermethylation of sperm DNA is not usually observed in either endogenous gene sequences¹¹ or in transgenic proviral sequences¹². The methylation status of the CAT 17 transgene in oocytes is not known. If it is methylated in oocytes, then it apparently survives in this form during embryogenesis. If, however, CAT 17 DNA were unmethylated in oocytes, then some other type of primary modification, for example protein binding to DNA, may interfere with *de novo* methylation of the paternal, but not the maternal copy. This type of interference has been suggested to account for methylation-free islands in the mouse genome¹³. Whatever the molecular basis of the chromosomal modifications, the observation of undermethylation in sperm and high somatic methylation in adult males carrying a maternally inherited transgene provides a basis for the determination of the precise timing of switching of the methylation pattern (see Fig. 2*c*) during male germ-cell differentiation.

Nuclear transplantation studies⁴ and genetic experiments¹⁴ show that maternal or paternal duplications of the genome or of particular chromosomal regions in the mouse produce phenotypes that manifest themselves late in embryogenesis. Thus some chromosomal regions must maintain imprints that distinguish parental alleles until late in development. Although there are known differences in the extent to which sperm and oocyte DNA is methylated, at least in repetitive DNA sequences¹⁵⁻¹⁷, it is not clear whether these differences persist upon replication

of chromosomes and whether they could mark maternal and paternal chromosomes in embryogenesis. Minor differences in methylation between the maternal and the paternal X chromosome in extraembryonic tissues¹⁸ appear to reflect the activity of the X chromosome and not necessarily its germline origin. Experiments are in progress to decide between the two possibilities using another transgenic strain with an X-linked integration, in which we have observed methylation differences in extraembryonic tissues (A.C. and W.R., unpublished results).

There is mounting evidence to suggest that genomic imprinting is of major importance in the control of mammalian embryogenesis¹⁹; it may also influence the phenotypic expression of genetic disorders in the human²⁰. Moreover, imprinting has recently been implicated in the regulation of mating type switching in yeast²¹. Genetic studies indicate that imprinting of particular chromosomal regions in the mouse may result in differential gene expression of parental alleles¹⁴. Although in strain CAT 17 the transgene was not expressed following either parental transmission (see Fig. 2 legend), we envisage that differential expression of transgenic markers will occur provided appropriate gene constructs are used. Some of these experiments are now in progress. But our observations suggest that the CAT-IgH-E DNA in the mouse strain CAT 17 is integrated into a region that is subject to chromosomal modification specific for the parental origin, and that methylation of the transgenic DNA reflects this property. Similar observations have recently been made in mice with different transgenes (C. Sapienza, personal communication; J. L. Swain, personal communication), for example an RSV-*myc* construct, in which methylation after maternal transmission and demethylation and expression following paternal transmission was observed²⁴. We do not know whether methylation is the primary imprinting signal or whether differential methylation is a consequence of primary modifications of a different nature. It is interesting to note that transgenic DNA can experience methylation changes that are induced by the chromosomal position of insertion, without concomitant changes of methylation in the surrounding cellular DNA¹². Thus, analysis of methylation of foreign DNA seems to be a useful marker for chromosomal modifications. Further analysis of the chromosomal region of integration in strain CAT 17 awaits molecular cloning of the cellular sequences flanking the foreign DNA. However, our results demonstrate that there are molecular mechanisms to mark homologous alleles in the mouse and that these modifications are germline-specific,

heritable and reversible. These are indeed the key features of chromosomal imprinting³.

We thank Michael Neuberger, Sarah Howlett and Nick Allen for help and discussions, Mr Bob for artwork and Linda Notton for typing the manuscript. W.R. was supported by a fellowship from EMBO and A.C. was supported by an AFRC studentship. We also thank the MRC Laboratory of Molecular Biology, Cambridge for support, and Dr Carmen Sapienza and Dr Judy Swain for communicating their findings to us before publication.

Received 16 April; accepted 26 May 1987.

1. Surani, M. A. H., Barton, S. C. & Norris, M. L. *Nature* **308**, 548-550 (1984).
2. McGrath, J. & Solter, D. *Cell* **37**, 179-183 (1984).
3. Surani, M. A. H., Reik, W., Norris, M. L. & Barton, S. C. *J. Embryol. exp. Morph.* **97** (Suppl.) 123-136 (1986).
4. Surani, M. A. H., Barton, S. C. & Norris, M. L. *Cell* **45**, 127-136 (1986).
5. Barton, S. C., Surani, M. A. H. & Norris, M. L. *Nature* **311**, 374-376 (1984).

6. Surani, M. A. H., Barton, S. C. & Norris, M. L. *Nature* **326**, 395-397 (1987).
7. Searle, A. G. & Beechey, C. V. in *Aneuploidy: Etiology and Mechanisms* (eds Dellarco, V. L., Voytyle, P. E. & Hollaender, A.) 363-376 (Plenum, New York, 1985).
8. Reik, W. *et al. Eur. J. Immun.* **17**, 465-469 (1987).
9. Chapman, V., Forrester, L., Sansford, J., Hastie, N. & Rossant, J. *Nature* **307**, 284-286 (1984).
10. Rossant, J., Sanford, J. P., Chapman, V. M. & Andrews, G. K. *Dev. Biol.* **117**, 567-573 (1986).
11. Rahe, B., Erickson, R. P. & Quinto, M. *Nucleic Acids Res.* **11**, 7947-7959 (1983).
12. Jahner, D. & Jaenisch, R. *Molec. cell. Biol.* **5**, 2212-2220 (1985).
13. Bird, A. P. *Nature* **321**, 209-213 (1986).
14. Cattanach, B. M. & Kirk, M. *Nature* **315**, 496-498 (1985).
15. Sanford, J., Forrester, W., Chapman, V., Chandley, A. & Hastie, N. *Nucleic Acids Res.* **12**, 2823-2936 (1984).
16. Ponzetto-Zimmerman, C. & Wolgemuth, D. J. *Nucleic Acids Res.* **12**, 2807-2822 (1984).
17. Monk, M., Boubelik, M. & Lehnert, S. *Development* **99**, 371-382 (1987).
18. Lock, L. F., Takagi, N. & Martin, G. R. *Cell* **48**, 39-46 (1987).
19. Surani, M. A. H. in *Experimental Approaches to Mammalian Embryonic Development* (eds Rossant, J. & Pedersen, R. A.) 401-435 (Cambridge University Press, 1986).
20. Myers, R. H., Madden, J. J., Teague, J. L. & Falek, A. *Am. J. Hum. Genet.* **34**, 481-488 (1982).
21. Klar, A. J. S. *Nature* **326**, 466-470 (1987).
22. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **137**, 266-267 (1984).
23. Hecht, N. B., Liem, H., Kleene, K. C., Distel, R. J. & Ho, S. M. *Dev. Biol.* **102**, 452-461 (1984).
24. Swain, J. L., Stewart, T. & Leder, P. *Clin. Res.* **35**, 419A (1987).

Degree of methylation of transgenes is dependent on gamete of origin

Carmen Sapienza*, Alan C. Peterson*, Janet Rossant† & Rudi Balling‡

* Ludwig Institute for Cancer Research (Montreal Branch), 687 Pine Avenue West, Montreal, Canada H3A 1A1

† Division of Molecular and Developmental Biology, Mount Sinai Hospital Research Institute, 600 University Avenue, Toronto, Canada M5G 1X5

Data derived from both pronuclear transplantation experiments and classical genetic experiments indicate that the maternal and paternal genetic contributions to the mammalian zygote nucleus do not function equivalently during subsequent development¹⁻⁵. These observations have been interpreted as resulting from differential 'genome imprinting' during male and female gametogenesis^{2,3}. The molecular mechanism responsible for genome imprinting is unknown, but data gathered to date require that the mechanism fulfil at least four criteria: (1) the imprint must be physically linked to the pronucleus^{2,3}; (2) the imprint must persist through DNA replication and cell division⁶; (3) the mechanism must be capable of affecting gene expression^{2,7}; and (4) the mechanism must be capable of switching the identity of the imprint from one sex to the other in successive generations^{2,3}. One molecular mechanism which could satisfy the first three criteria is differential DNA methylation during gametogenesis itself^{6,8}, or before formation of the zygote nucleus during embryogenesis. We present data indicating that the methylation patterns of exogenous DNA sequences in transgenic mice can be changed by switching their gamete of origin in successive generations. These data suggest that DNA methylation can also satisfy the fourth criterion for an imprinting mechanism.

‡ Present address: Department of Molecular Cell Biology, Max Planck Institute of Biophysical Chemistry, Postfach 2841, 3400 Göttingen, West Germany.

Examination of the methylation states of some endogenous repetitive⁹⁻¹¹ and structural gene sequences (J. B. Sanford, J.R. and V. M. Chapman, unpublished data) has revealed differences between oocyte and sperm DNA, consistent with a role for DNA methylation in imprinting. But the presence of two alleles (one maternally derived and one paternally derived) at most autosomal loci normally precludes the direct examination of such loci for a 'methylation imprint'. Introduction of new genetic information into the mouse genome by DNA microinjection creates new genetic loci. Individuals carrying such loci can be maintained as hemizygotes, and transmit the newly-acquired transgene locus to only 50% of their progeny. These characteristics allow one to control the gametogenic history of these loci, and assess whether their methylation state is affected by that history.

Preliminary observations by one of us (R.B.) on three transgenic lines containing a *Drosophila hsp70-Escherichia coli lac Z* construct suggested that, in one line, some *MspI* sites within the *lac Z* gene were less methylated when inherited from a male than from a female. Unfortunately, the informative line became extinct. We therefore continued this investigation by examining the methylation status of the transgene loci in five different lines which are hemizygous for the quail fast fibre muscle-specific troponin I (TNI) gene (ref. 12, P. Hallauer *et al.* unpublished data). The pedigree of one of these is shown in Fig. 1.

The transgene methylation pattern was determined by Southern blotting of tail DNAs digested with methylation-sensitive restriction endonucleases. When these DNAs were digested with *Bam*HI and *Hpa*II and the blots hybridized with a *Bam*HI-*Kpn*I fragment of the TNI gene, the methylation state of several *MspI* sites within the 3' portion of the TNI gene could be determined (Fig. 2a).

In four of the five lines, we obtained evidence that the methylation state of the transgene varied according to the gamete of origin. The most extensive analyses were carried out on line 379 (Fig. 1) which contains ~20 copies of the TNI gene inserted as a single tandem array. In this line, the transgene consistently became more methylated when inherited from a female parent,

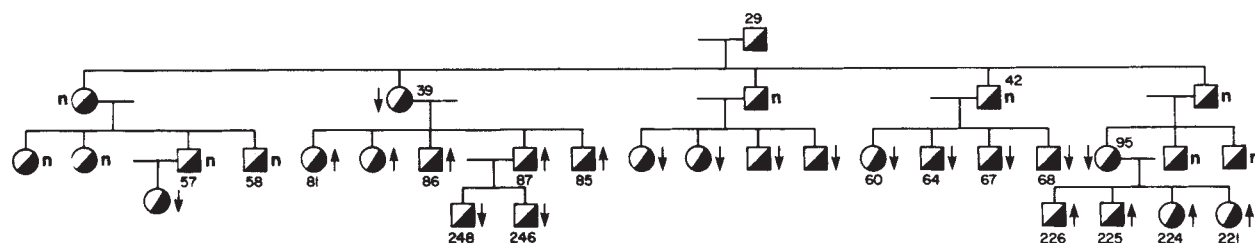


Fig. 1 Pedigree of transgenic line 379. All individuals shown in the pedigree are hemizygous for the transgene locus, consisting of ~20 copies of the quail fast-fibre type isoform of the troponin I gene (ref. 12 and Fig. 2a). All matings were to non-transgenic C57BL/6J, B6D2F1 or B6C3F1 mice. Numbers refer to individuals shown in Figs 2-4; arrows, direction of the change in transgene methylation state; 'n', no change in transgene methylation from the previous generation (see text).

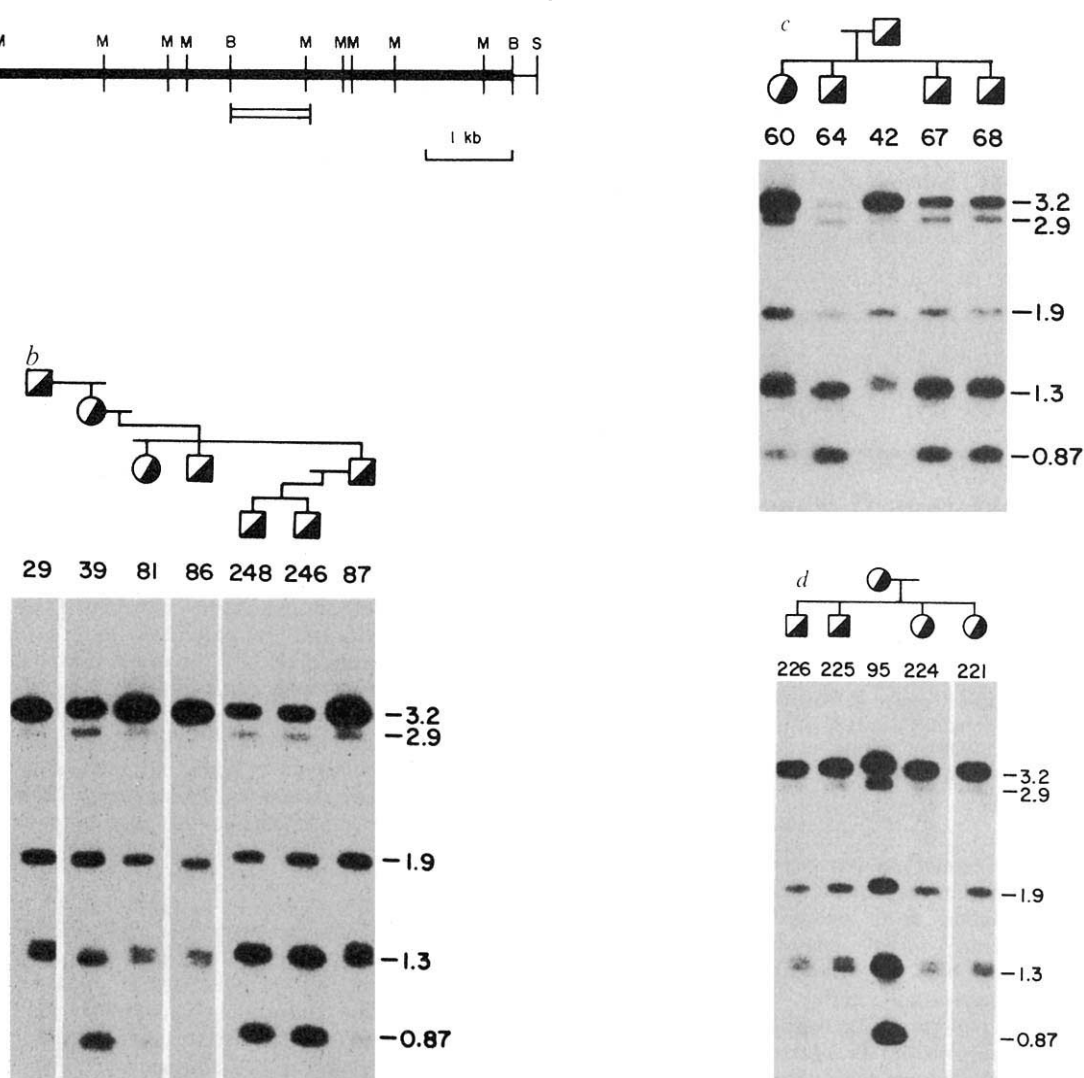


Fig. 2. Detailed excerpts of Fig. 1. High-molecular-weight genomic DNAs were prepared from tail biopsies¹⁹. *a*, Restriction endonuclease cleavage map of the quail fast-fibre type troponin I gene¹². E, *Eco*RI; B, *Bam*HI; M, *Msp*I; S, *Sal*I. The entire 7.1-kb *Eco*RI-*Sal*I fragment was injected into pronuclear stage mouse zygotes¹⁹. Thin horizontal line, pBR322 sequences; thick horizontal line, quail DNA containing the entire troponin I coding sequence, as well as ~600 bp of 5' and 1,200 bp of 3'-flanking sequence. Open box, location of the 880-bp *Bam*HI-*Kpn*I fragment used as a hybridization probe in Figs 2*b-d*, 3*a* and 4. *b*, Individuals comprising four generations of line 379. The methylation state of the transgene decreases when inherited from a male (female 39 and males 246 and 248) and increases when inherited from a female (female 81 and males 86 and 87). *c*, Male 42 and offspring, demonstrating a reduction in the methylation state of the transgene following inheritance through a male; *d*, female 95 and offspring, demonstrating an increase in the methylation state of the transgene following inheritance through a female.

Methods. Aliquots of 5 µg of each DNA were digested with *Bam*HI (72U; Pharmacia) and *Hpa*II (54U; Pharmacia); the resulting fragments were resolved by electrophoresis on 0.75% agarose gels²⁰, transferred to a nylon membrane (Nytran) and hybridized with a ³²P-labelled²⁰ 880-bp probe (encompassing parts of exons V and VII, all of exon VI and all of introns 5 and 6¹²). After washing (0.5×SSC, 0.1% SDS, 65 °C) the membranes were exposed to X-ray film (Fuji RX) at 80 °C (Lightning Plus intensifying screens, Dupont) for 16–72 h. Completion of restriction endonuclease digestion was monitored in all cases by the addition of a control plasmid DNA to an aliquot of the experimental sample digest²¹. A reaction was judged to be complete when the internal control gave a digestion pattern identical to plasmid DNA alone cleaved with the same restriction endonuclease.

and in some offspring, less methylated when inherited from a male parent. Genomic blots of selected members of the 379 pedigree are shown in Fig. 2*b-d*. All members of the pedigree contain the 3.2-kilobase (kb) *Bam*HI fragment, indicating that some copies of the transgene are completely methylated over this interval. But lower molecular weight bands indicative of digestion at internal *Msp*I sites were also observed. The relative intensities of the 3.2-kb band and the lower molecular weight bands provide a measure of the methylation state of the transgene in different mice. For example, in Fig. 2*b*, female 39 shows

a decrease in the intensity of the 3.2-kb band and a clear increase in the intensity of the 870-base-pair (bp) band when compared with her father (male 29). Her offspring (female 81, males 86 and 87) show an increase in methylation as measured by a sharp reduction in the intensity of the 870-bp band and increase in the intensity of the 3.2-kb band. Offspring of male 87 show a decrease in methylation, producing a pattern similar to female 39. Similar gamete-of-origin-dependent shifts are shown in Fig. 2*c, d*. Analyses of genomic DNA blots of these same mice with probes derived from different regions of the gene, or the entire

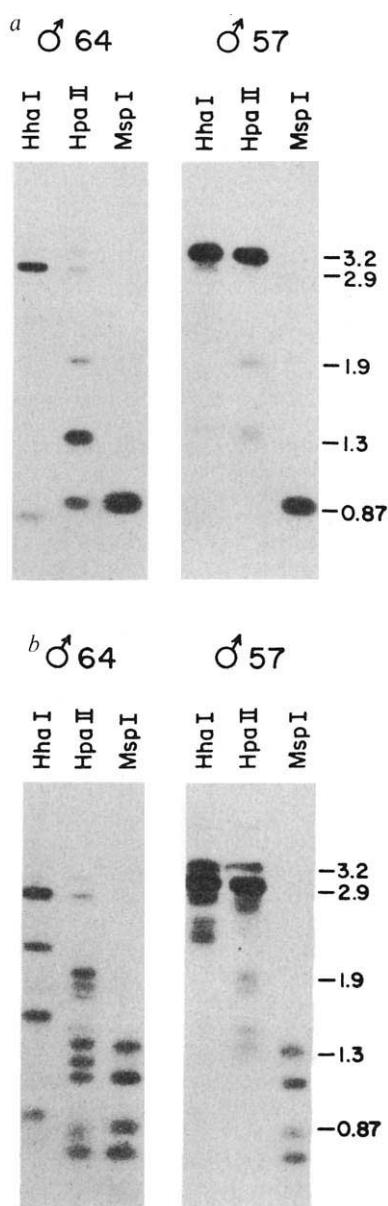


Fig. 3 *a*, Tail DNAs from males 64 (transgene inherited from a male) and 57 (transgene inherited from a female) digested with *Bam*HI, followed by digestion with the indicated restriction endonuclease and hybridized with the 880-pb probe (see Fig. 2*a*). *b*, Same blot as in *a*, hybridized with the entire 7.1-kb *Eco*RI-*Sal*I fragment (Fig. 2*a*).

gene, gave qualitatively similar results (Fig. 3*b* and data not shown). Furthermore, gametic changes in methylation of the transgene are not restricted to *Msp*I sites. Qualitatively similar results are also obtained with the methylation-sensitive enzyme *Hha*I (Fig. 3), as well as several others (data not shown).

It is clear from this analysis that the methylation state of some restriction-endonuclease cleavage sites within the transgene locus carried by line 379 can vary with male or female germ line transmission in a manner consistent with the requirements for a genomic imprinting process. Evidence that at least some of the changes in methylation imposed during germline transmission actually occur during gametogenesis is provided by a comparison of the methylation pattern of DNA from testes and somatic tissues of males which have inherited the transgene from a female (Fig. 4). Testes DNA is clearly undermethylated

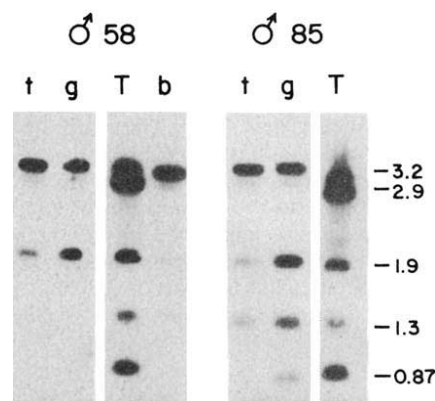


Fig. 4 DNAs from tissues of males 58 and 85 (transgene inherited from a female) digested with *Bam*HI and *Hpa*II and hybridized with the 880-bp probe. t, Tail; g, gastrocnemius; b, brain; T, testes.

when compared with DNA from tail, as well as other somatic tissues analysed (brain, liver, kidney — Fig. 4 and data not shown). The only somatic tissue DNAs analysed which showed a difference in the methylation state of the TNI transgene were obtained from striated muscle (Fig. 4 and data not shown).

Within this transgenic line, males 57 and 58, as well as their female siblings, did not show an increase in the methylation state of the transgene. Because their mother already exhibited a transgene methylation pattern similar to that observed for individuals which had inherited the transgene from a female (81, 85, 86 and 87 as well as 221, 224, 225 and 226) such a change was not expected. But some offspring of some males failed to show the anticipated decrease in transgene methylation state. The significance of this variation is unclear. The differences could result from activities expressed post fertilization¹³, but they could also reflect gametogenic events requiring that all products of meiosis do not share an identical methylation status¹⁴.

Less extensive analyses of the other four transgenic lines indicated that two show evidence for changes similar to those observed in line 379, one shows no effect and one shows evidence for changes in the reverse direction, that is, sequences become more methylated after transmission through the male germ line (data not shown).

Our combined results indicate that a DNA methylation imprint imposed during gametogenesis can persist into the adult and can switch identity in successive generations. Because the same transgene can show different methylation states and switching behaviour in different lines, it seems likely that the observed methylation changes reflect events influenced by flanking sequences in the host genome. Two models for the role of gametic DNA methylation in genome imprinting can be envisaged. In one, the only differences that persist through development are those that occur in genes that show obligate differential maternal/paternal expression⁷. In the other model, methylation differences between male and female derived genetic information are widespread, but only some genes respond by exhibiting differential expression. The latter model is consistent with the relatively high frequency of observed gametic methylation switches in different transgenic lines, and with known variable effects of DNA methylation on gene expression¹⁵. Gametic methylation differences imposed on transgenes as a result of their integration into imprinted regions of the host genome might affect their expression. But if the second model is correct, such effects may occur independently of whether flanking sequences show differential maternal/paternal expression.

Should the operation of such a methylation-based imprinting mechanism, described here for transgenes, also operate on

endogenous loci, it could account for the observed quantitative variations in phenotype which are not reflected in genotype^{14,16-18}.

We thank Jean Paquette and Thu Hang Tran for technical assistance, Pat Hallauer and Irene Tretjakoff for transgenic mouse production and Sue Varmuza for stimulating discussion. We also thank Azim Surani and members of his laboratory for stimulating discussion and communication of results prior to publication. This work was supported in part by grants from the NCI (to J.R.), NSERC (to J.R.) and MRC (to J.R.) of Canada. J.R. is a Research Associate of the NCI of Canada. R.B. was supported by a fellowship from the Deutsche Forschungsgemeinschaft.

Received 30 March; accepted 15 June 1987.

- McGrath, J. & Solter, D. *Science* **220**, 1034-1036 (1983).
- Surani, M. A. H., Barton, S. C. & Norris, M. L. *Nature* **308**, 548-550 (1984).
- McGrath, J. & Solter, D. *Cell* **37**, 179-183 (1984).
- Cattanach, B. M. & Kirk, M. *Nature* **315**, 496-498 (1985).
- Surani, M. A. H., Barton, S. C. & Norris, M. L. *Nature* **326**, 395-397 (1987).
- Surani, M. A. H., Barton, S. C. & Norris, M. L. *Cell* **45**, 127-136 (1986).
- McGrath, J. & Solter, D. *Nature* **308**, 550-551 (1984).
- Surani, M. A. H. in *Experimental Approaches to Mammalian Embryonic Development* (eds Rossant, J. & Pedersen, R. A.) 401-435 (Cambridge University Press, 1986).
- Sanford, J., Forrester, L., Chapman, V. M., Chandley, A. & Hastie, N. *Nucleic Acids. Res.* **12**, 2823-2836 (1984).
- Sanford, J. P., Chapman, V. M. & Rossant, J. *Trends Genet.* **1**, 89-93 (1985).
- Monk, M., Boubelik, M. & Lehnert, S. *Development* **99**, 371-382 (1987).
- Baldwin, A. S., Kittler, E. & Emerson, C. P., Jr *Proc. natn. Acad. Sci. U.S.A.* **82**, 8080-8084 (1985).
- Jaehner, D. *et al. Nature* **298**, 623-628 (1982).
- Klar, A. J. S. *Nature* **326**, 466-470 (1987).
- Bird, A. P. *Nature* **321**, 209-213 (1986).
- Myers, R. H. *et al. Lancet* **i**, 208-210 (1983).
- Farrer, L. A. & Connealy, P. M. *Am. J. hum. Genet.* **37**, 350-357 (1985).
- Froster-Iskenius, U., McGillivray, B. C., Dill, F. J., Hall, J. G. & Herbst, D. S. *Am. J. Med. Genet.* **23**, 619-631 (1986).
- Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo, A Laboratory Manual* (Cold Spring Harbor Laboratory, 1986).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, 1982).
- Wyman, A. R. & White, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6754-6758 (1980).

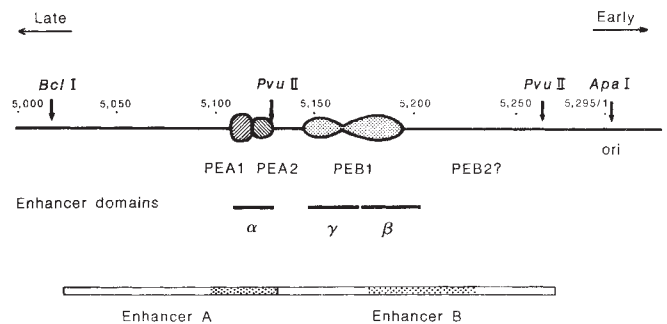


Fig. 1 Functional map and principal features of the Py enhancer. Horizontal arrows, directions of late and early transcription; vertical arrows, restriction sites for *Bcl*I, *Pvu*II and *Apa*I; 'ori', origin of viral DNA replication. Also shown diagrammatically are enhancer-protein interactions of nuclear factors from mouse 3T6 cells: PEA1 (with nucleotide positions 5,113-5,123), PEA2 (positions 5,120-5,128), PEB1 (positions 5,146-5,196)^{9,10} and PEB2 whose interaction with the origin-proximal region of enhancer B is currently under investigation. The three functional α -, β - and γ -domains, defined by Veldman *et al.*³ and Hassel *et al.*⁴ are represented by lines. Open boxes, the two enhancers A and B described by Herbomel *et al.*²; dotted areas, core sequences.

of viral gene expression during cellular differentiation, and perhaps more generally in the control of gene expression during early embryonic development.

We have shown recently, using gel retardation and DNase I footprinting assays, that nuclear proteins from mouse 3T6 cells interact with the α -, β - and γ -domains of the Py enhancer^{5,9,10} (Fig. 1). The α - and β -domains are the cores of the A and B enhancers respectively; the γ -domain is an auxiliary element of the B enhancer²⁻⁴. A single factor (PEB1) footprints on the β - γ -domain and two factors (PEA1 and PEA2) footprint on the α -domain (nucleotide position 5,109-5,130) of the enhancer. Several repeats of the α -domain, or a single copy of the β - γ -domain constitute the minimal sequences required for enhancer function^{3,4}, suggesting that these DNA-protein interactions must be crucial for the biological activity of the virus.

As the block to polyoma virus growth in EC cells seems to involve at least in part a deficient A enhancer^{2,13}, we wanted to check whether the previously characterized factors binding to this enhancer domain were present in undifferentiated cells. Nuclear extracts from F9 cells were tested for their ability to interact with the enhancer sequences of polyoma virus in conditions identical to those used with 3T6 cell extracts. The DNase I footprint obtained is shown in Fig. 2. 3T6 and F9 cell extracts display a rather similar footprint on the β - γ -domain, whereas no interaction can be detected on the α -domain with the F9 cell extract. Proteins of F9 cells give a more complete protection of the β - γ -domain than proteins of 3T6 cells, but gel retardation assays showed that both extracts form the same type of complex (data not shown). Some differences in the DNase I digestion patterns can also be observed in the origin-proximal part of the B enhancer, but these features are currently under further investigation and will be described elsewhere (factor PEB2 in Fig. 1). Nuclear extracts from EC PCC4 cells were tested in the same conditions and gave exactly the same footprint as F9 cell extracts on the Py enhancer (data not shown). PEA1 and PEA2 are thus inactive or absent in nuclear extracts from EC cells, or the α -binding activity is reduced to such a low level that we cannot detect it in our footprinting assay. This does not seem to result from an inability to extract nuclear factors from EC cells, because we detect abundant PEB1 activity in these extracts. Furthermore, it is not a result of loss of factors during preparation of the nuclei, because whole cell extracts prepared from F9 cells gave the same results (data not shown).

Induction of a factor that binds to the polyoma virus A enhancer on differentiation of embryonal carcinoma cells

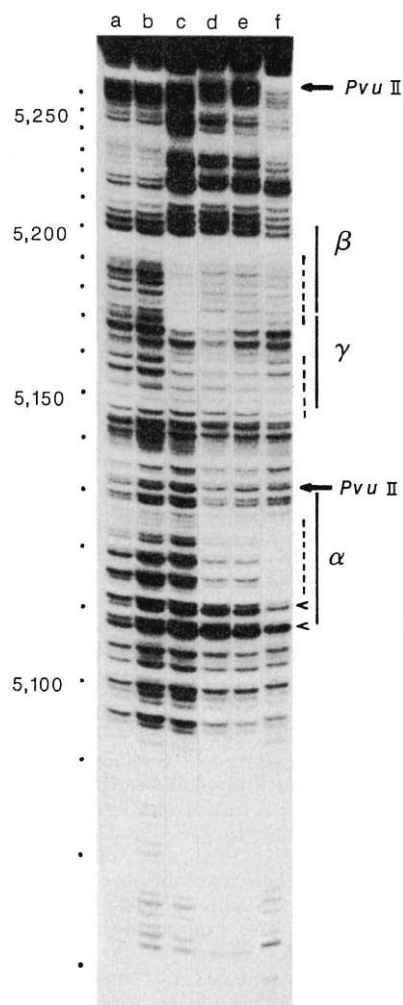
Marie-Hélène Kryszke, Jacques Piette & Moshe Yaniv

UA 041149 CNRS, Département de Biologie Moléculaire, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France

The cell type specificity of certain enhancers, competition experiments and the interactions that occur between proteins and enhancer sequences demonstrate that enhancers are the targets of specific factors involved in transcription control¹. The 246-base pair *Bcl*I-*Pvu*II restriction enzyme fragment of polyoma virus has been shown to include two distinct enhancers, Py A and Py B, composed of several subdomains²⁻⁴ which interact with nuclear proteins from mouse fibroblasts⁵⁻¹⁰. Embryonal carcinoma (EC) cells do not permit polyoma virus infection: both viral transcription and DNA replication are blocked¹¹. Host-range mutants of polyoma virus (EC mutants) capable of overcoming the expression block in EC cells have mutations or sequence rearrangements in their enhancer region¹². In an attempt to understand the molecular basis of this host restriction we compared the binding patterns displayed on the viral enhancer sequences by nuclear proteins prepared from EC cells or from fibroblasts. We show that one of the fibroblast factors required for Py A enhancer function, almost undetectable in EC cells, is induced after differentiation of these cells into parietal endoderm, suggesting that this protein is crucial in the regulation

Fig. 2 Differentiation-stage-specific footprints on the Py enhancer. DNase I digestion patterns are shown here with the protected sequences (dotted lines), hypersensitive sites (ν), and the α -, β - and γ -domains indicated alongside. Chemical degradation products of the end-labelled fragment run in parallel to the DNase I digestion products were used to localize the DNase I cleavage sites relative to the nucleotide sequence (not shown, see ref. 10). Lanes a and b, DNA treated with DNase I in the absence of extract. Lanes c-f, same DNA with extract as follows: lane c, undifferentiated F9 cells; lane d, F9 cells treated with RA alone (2×10^{-7} M); lane e, F9 cells treated with RA (2×10^{-7} M) and DB-cAMP (10^{-3} M), lane f, 3T6 cells.

Methods. Cells were grown in Dulbecco's modified Eagle's medium supplemented with 7% (3T6) or 15% (F9) fetal calf serum. Nuclei from 3T6 cells were prepared as described previously⁵. F9 nuclei were obtained by hypotonic lysis of the cells in RSB buffer (10 mM Tris-HCl, pH 7.5; 10 mM NaCl; 3 mM MgCl₂). Nuclear extracts from the different cell types were prepared by washing isolated nuclei with 0.4 M NaCl (for details, see ref. 5). Protein (20 μ g) from each extract was incubated on ice with 150 ng poly(dI-dC) in 10 mM HEPES, pH 8; 17.5% glycerol; 0.1 mM EDTA; 20 mM KCl; 4 mM MgCl₂; 2 mM dithiothreitol (DTT); 4 mM spermidine in a total volume of 10 μ l. After 10 min, the radioactively end-labelled DNA fragment containing the Py enhancer (*Bcl*I-*Ap*al fragment) was added and the incubation continued for another 10 min on ice. The mixture was then treated with 2 μ l DNase I (100-400 μ g ml⁻¹) for 1 min at 20 °C. The reaction was stopped with 12 μ l 0.1% SDS; 50 mM EDTA; 200 μ g ml⁻¹ transfer RNA. The DNA was purified by proteinase-K digestion, phenol extraction and ethanol precipitation, and was analysed on a sequencing gel.



F9 cells become permissive to polyoma virus when they differentiate into parietal endoderm after treatment with retinoic acid (RA) in the presence or absence of dibutyryl-cAMP (DB-cAMP)¹⁴⁻¹⁶. We performed DNase I footprinting experiments with nuclear extracts from F9 cells treated with RA for 7 days, or RA with DB-cAMP for 5 days. The footprints obtained in both cases resemble that given by 3T6 cell extracts: two hypersensitive sites (at positions 5,108 and 5,111) appear on the late side of the α -domain together with a partial protection of this domain (Fig. 2). Thus, the α -binding activity becomes detectable after differentiation of F9 cells.

To support these results and to confirm the identity of the factors detected after differentiation, the more sensitive gel retardation technique was used. PEA1 interacts with several viral and cellular enhancers including SV40, *c-fos*, H-2K and so on (ref. 10 and J.P., unpublished results). To check for the presence of this factor in our extracts we used a specific probe, the PEA1 binding site located in the *fos* enhancer, with one base pair mutated to generate an ideal palindrome displaying a stronger interaction with the protein (ref. 10; J.P., unpublished observations). The same protein-DNA complex is formed in the presence of differentiated F9 cell extracts and in the presence of 3T6 cell extracts (Fig. 3A). Very small amounts of this complex can be seen with the F9 cell extracts. In each case it can be specifically titrated by the homologous sequence (competitor oligonucleotide PK, lanes d) as well as by the oligonucleotide A2 (lanes c) which binds PEA1 but fails to be recognized by PEA2 (ref. 10), whereas excess oligonucleotide A1, capable of binding PEA2 only¹⁰, does not interfere with complex formation. (See Fig. 3B for the oligonucleotides used.) No PEA2 activity

could be detected either in undifferentiated or in differentiated F9 cells by gel retardation assays (data not shown). Thus we clearly demonstrate that PEA1, almost undetectable in undifferentiated F9 cells, is expressed after differentiation of these cells. However, it seems to be less abundant or less active in the endoderm-like cells than in fibroblasts.

These results are in very good agreement with those from transient and stable expression assays performed in differentiated and EC cells^{2,17}. Although the B enhancer has the same low efficiency in 3T6 and EC cells, the A enhancer, the major one in 3T6 cells, is 3-4-fold less active in EC cells than in fibroblasts. The limiting level of PEA1 in EC cells could account for the weakness of the A enhancer in this context. Wasylyk *et al.*¹⁸ have shown that the Py enhancer can be activated by the *c-Ha-ras* oncogene in F9 cells but not in fibroblasts and that in differentiated F9 cells an intermediate stimulation was obtained. Depending on the respective pre-existing level of PEA1 activity in the different cell types, the product of the *c-ras* gene could induce a varying additional stimulation of the A enhancer by further increasing the synthesis or the activity of this factor. In addition to its role in transcription, the Py enhancer *cis* activates viral DNA replication *in vivo*⁴. The α -domain together with the core of the replication origin is sufficient to allow Py replication in differentiated cells but not in two-cell embryos¹³. Again our results could provide an explanation for this differentiation-stage-specific regulation of the A enhancer activity. One of the polyoma virus mutants defective for viral replication and transcription¹⁹, B1, bears several point mutations in its α -domain which prevent this sequence from binding PEA1 *in vitro*. Reversion of one of these mutations (revertant B1-5140) was shown

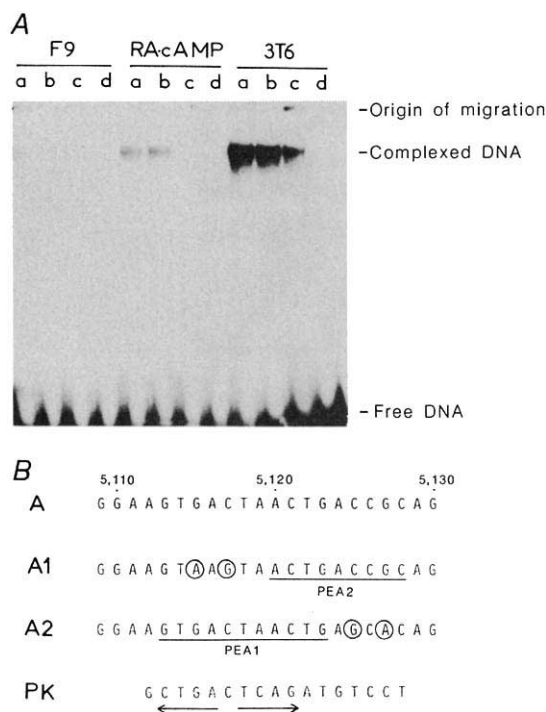


Fig. 3 Gel retardation assay. **A**, Autoradiography of the gel. The first four lanes correspond to undifferentiated F9 cells (F9), the next four lanes to F9 cells treated with RA and DB-cAMP (RA, cAMP) and the last four to 3T6 cells (3T6). The competitor oligonucleotides used were C (lanes a), A1 (lanes b), A2 (lanes c) and PK (lanes d). **B**, Sequences of the oligonucleotides used. The Py A wild-type sequence corresponding to the α -domain is sequence A. A1 is the same sequence with two point mutations (positions 5,115 and 5,117) that abolish PEA1 binding¹⁰. The mutated nucleotides are pointed out by circles. A2 bears two point mutations (positions 5,125 and 5,127) which prevent the binding of PEA2 (ref. 10). The binding sites for PEA2 and PEA1 are represented under sequences A1 and A2, respectively. PK is the oligonucleotide described above with its palindromic sequence indicated by horizontal arrows.

Methods. The incubation was performed as in the Fig. 2 legend. Protein (12 μ g) from F9 cells and (6 μ g) from 3T6 cells were used. Unlabelled competitor oligonucleotide (2.5 pmol) (C, corresponding to the Py γ -domain and used as nonspecific competitor, A1, A2 or PK—see **B**) were added to the preincubation mixture. The radioactively labelled PK probe is a double-stranded oligonucleotide containing a perfectly palindromic PEA1 binding site derived from the *c-fos* sequence (positions -307 to -291, ref. 28). Subsequently 10 μ l of NM buffer (10 mM HEPES pH 8; 17.5% glycerol; 0.1 mM EDTA; 100 mM NaCl; 10 mM MgCl₂; 2 mM DTT; 100 μ g ml⁻¹ bovine serum albumin (BSA), containing 0.1 mg ml⁻¹ sonicated salmon sperm DNA were added to the samples just before loading on an 8% native polyacrylamide gel. After electrophoresis at 160 V, the gel was dried and autoradiographed.

to restore both PEA1 binding *in vitro* and viral expression *in vivo* (J.P. and W. R. Folk, unpublished results). These observations strongly support the hypothesis that PEA1 is important in the biology of the virus. At present we cannot decide whether the rate of synthesis of PEA1 dramatically increases at a specific developmental stage, or whether this factor preexists in an inactive form and is then modified to become capable of efficient binding to polyoma virus DNA. In the case of the κ -light-chain enhancer binding factor NF- κ B, a post-translational mechanism has been postulated to account for the specific induction of the enhancer-binding activity during B cell differentiation²⁰.

The lack of activity of the A enhancer in EC cells is known to be compensated for in two ways. Polyoma virus EC PCC4

mutants duplicate their α -domain, increasing the affinity of the limiting cellular factors for the enhancer, whereas EC F9 mutants create a novel, stronger enhancer motif by mutation in the B enhancer². In F9 cells an additional feature must render the block to polyoma virus replication more stringent than in PCC4 cells because EC PCC4 mutants cannot develop in F9 cells whereas EC F9 mutants grow on both cell types¹².

Other viruses, like simian virus 40 (SV40) and type C retroviruses, unable to grow in EC cells, also show reduced activity of their enhancers in these cells, suggesting the possibility of a common developmental regulation of these elements^{21,22}. Note that the α -domain of the Py enhancer displays sequence similarity to the SV40 enhancer, that PEA1 also binds to the SV40 enhancer¹⁰, and that both enhancers are repressed by the E1A products²³. Factor PEA1 is probably identical to protein AP1 described by Lee *et al.*²⁴. An additional interesting feature is that PEA1 also binds to the enhancer-like sequences of the mouse MHC gene *H-2K* (J.P., unpublished results), which is developmentally regulated and expressed only after differentiation of F9 cells²⁵.

Because F9 cells contain an E1A-like activity which disappears following differentiation²⁶, and Py EC F9-1 enhancer is insensitive to repression by E1A (ref. 27), it has been suggested that the wild-type Py enhancer would be blocked in EC cells by the binding of an E1A-dependent negative factor. On the contrary, the fact that hybrid cells obtained by fusion between EC cells and permissive cells are able to support viral growth¹¹, and our present results, suggest that the expression block in EC cells might be due to the absence of a positive factor affecting enhancer activity. Moreover, the wild-type and F9-1 mutated Py enhancers are equally stimulated by *ras* in F9 cells, suggesting that *ras* and the F9-1 mutation activate the enhancer by two independent mechanisms¹⁸. At present, the functional complexity of the Py enhancer makes it seem likely that more than one event is involved in the differentiation stage-specific control of viral expression. Purification of the multiple factors interacting with the different domains of the enhancer should help in answering these questions.

We thank M. Weiss and P. Herbolme for comments on the manuscript; M. Karin, R. Tjian, B. Wasyluk and W. R. Folk for communication of unpublished results. This work was supported by grants from the Centre National de la Recherche Scientifique and the Association pour la Recherche sur le Cancer.

Received 27 April; accepted 22 May 1987.

- Serfling, E., Jasin, M. & Schaffner, W. *Trends Genet.* **1**, 224-230 (1985).
- Herbolme, P., Bourachot, B. & Yaniv, M. *Cell* **39**, 653-662 (1984).
- Veldman, G. M., Lupton, J. & Kamen, R. *Molec. cell. Biol.* **5**, 649-658 (1985).
- Hassel, J. A., Muller, W. J. & Mueller, C. R. in *Cancer Cells* Vol. 4 (eds Botchan, M., Grodzicker, T. & Sharp, P.) 103-113 (Cold Spring Harbor Laboratory, New York, 1986).
- Piette, J., Kryszke, M.-H. & Yaniv, M. *EMBO J.* **4**, 2675-2685 (1985).
- Böhnlein, E. & Gruss, P. *Molec. cell. Biol.* **6**, 1401-1411 (1986).
- Fujimura, F. K. *Nucleic Acids Res.* **14**, 2845-2863 (1986).
- Ostapchuk, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 8550-8554 (1986).
- Piette, J. & Yaniv, M. *Nucleic Acids Res.* **14**, 9595-9611 (1986).
- Piette, J. & Yaniv, M. *EMBO J.* **6**, 1331-1337 (1987).
- Kelly, F. & Condamine, H. *Biochim. biophys. Acta* **651**, 105-141 (1982).
- Amati, P. *Cell* **43**, 561-562 (1985).
- Wirak, D. O. *et al. Molec. cell. Biol.* **5**, 2924-2935 (1985).
- Strickland, S. & Mahdavi, V. *Cell* **15**, 393-403 (1978).
- Strickland, S., Smith, K. K. & Marotti, K. R. *Cell* **21**, 347-355 (1980).
- Fujimura, F. K., Silbert, P. E., Eckhart, W. & Linney, E. *J. Virol.* **39**, 306-312 (1981).
- Linney, E. & Donnelly, S. *Cell* **35**, 693-699 (1983).
- Wasyluk, C., Imter, J. L., Perez-Mutul, J. & Wasyluk, B. *Cell* **48**, 525-534 (1987).
- Tang, W. J., Berger, S. L., Triesenberg, S. J. & Folk, W. R. *Molec. cell. Biol.* **7**, 1681-1690 (1987).
- Sen, R. & Baltimore, D. *Cell* **47**, 921-928 (1986).
- Herbolme, P., De Crombrughe, B. & Yaniv, M. in *Teratocarcinoma Stem Cells* (eds Silver, L. M., Martin, G. R. & Strickland, S.) 285-294 (Cold Spring Harbor Laboratory, New York, 1983).
- Linney, E., Davis, B., Overhauser, J., Chao, E. & Fan, H. *Nature* **308**, 470-472 (1984).
- Borrelli, E., Hen, R. & Chambon, P. *Nature* **312**, 608-612 (1984).
- Lee, W., Haslinger, A., Karin, M. & Tjian, R. *Nature* **325**, 368-372 (1987).
- Gachelin, G. *Biochim. biophys. Acta* **516**, 27-60 (1978).
- Imperiale, M., Kao, H. T., Feldman, L. T., Nevins, J. R. & Strickland, S. *Molec. cell. Biol.* **4**, 867-874 (1984).
- Hen, R., Borrelli, E., Fromental, C., Sassone-Corsi, P. & Chambon, P. *Nature* **312**, 249-251 (1986).
- Treisman, R. *Cell* **42**, 889-902 (1985).

Safety and immunogenicity in man of a synthetic peptide malaria vaccine against *Plasmodium falciparum* sporozoites

Deirdre A. Herrington*, David F. Clyde*, Genevieve Losonsky*, Manuel Cortesia*, James R. Murphy*, Jonathan Davis*, Shahida Baqar*, Arthur M. Felix†, Edgar P. Heimer†, Dieter Gillesen‡, Elizabeth Nardin§, Ruth S. Nussenzeig #, Victor Nussenzeig #, Michael R. Hollingdale¶ & Myron M. Levine*

*Geographical Medicine Division, Department of Medicine, University of Maryland School of Medicine and Center for Vaccine Development, Baltimore, Maryland 21201, USA

†Peptide Research Department, Hoffman-La Roche Inc., Nutley, New Jersey 07110, USA

‡Central Research Department, F. Hoffman-La Roche and Co., Ltd, Basle, Switzerland

§Department of Medical and Molecular Parasitology and

Department of Pathology, New York University School of Medicine, 550 First Avenue, New York, 10016, USA

¶Biomedical Research Institute, 12111 Parklawn Drive, Rockville, Maryland 20852, USA

A 12 amino-acid synthetic peptide (NANP)₃ comprising the immunodominant epitope of *Plasmodium falciparum* circumsporozoite protein was conjugated to tetanus toxoid (TT), adjuvanted with aluminium hydroxide, and administered intramuscularly in three doses at monthly intervals to 35 healthy males as a malaria vaccine. No significant adverse reactions were noted, with mild soreness at the injection site the only common symptom. Seroconversions against NANP occurred in 53% and 71% of recipients of 100 or 160 µg, respectively, measured by enzyme-linked immunosorbent assay (ELISA). Most ELISA-positive sera reacted with sporozoites by indirect immunofluorescence (IFA). Three vaccinees with the highest ELISA and IFA titres and four unimmunized controls were challenged with *P. falciparum* sporozoites introduced via the bites of infective *Anopheles* mosquitoes. Blood stage parasites were detected in all controls by 10 days (mean 8.5 days, range 7–10). In contrast, the two vaccinees who became infected did not manifest parasitaemia until day 11 and the third vaccinee showed neither parasites nor symptoms during the 29 day observation period. This first synthetic peptide parenteral vaccine against a communicable disease tested in man is safe and stimulates biologically active antibodies. These observations encourage the development of improved vaccine formulations which, by enhancing immunogenicity, may lead to practical vaccines to assist in the control of falciparum malaria.

Plasmodium falciparum malaria is undergoing a worldwide resurgence because of the spread of drug resistant parasite strains and increasing resistance of mosquitoes to insecticides¹. Malaria vaccines are being developed that may complement traditional malaria control measures².

In the early 1970s, complete protection against malaria in small numbers of volunteers was achieved by 'vaccination' with attenuated, X-irradiated sporozoites^{3–5} and was correlated with antibodies directed against the circumsporozoite (CS) protein which covers the surface of the mature sporozoite⁶. Recognition that the immunodominant epitope of the *P. falciparum* CS protein is a repetitive four amino-acid sequence (asparagine-alanine-asparagine-proline, or NANP)^{7–9} and that monoclonal antibodies directed against this sequence neutralize sporozoite infectivity *in vitro*¹⁰, paved the way for the development of sporozoite vaccine candidates prepared by synthetic peptide and recombinant DNA techniques^{2,10,11}. Synthetic NANP pep-

Table 1 Vaccination schedule and dosages of (NANP)₃-TT

Vaccine dose (µg)	No. of vaccinees	No. of controls
20	3	1
50	3	1
100	15	4
160	14	4

Immunizations were given into the triceps muscle on days 0, 28 and 56. Three volunteers in the 160 µg vaccine group received a third booster injection on day 133 in preparation for participation in the vaccine efficacy study. Ac-Cys (NANP)₃-OH was prepared by solid-phase peptide synthesis starting with proline-hydroxymethyl-resin (21.0 g, 0.36 nmol g⁻¹, 7.56 mmol) and 12 cycles of coupling reactions using 4 equivalents of *N*-α-*t*-butyloxycarbonyl (Boc)-amino acids with dicyclohexylcarbodiimide (DCC) couplings. Boc-asparagine was introduced as the *N*-hydroxybenzotriazole-ester. The final cycle was carried out with Boc-S-3,4-dimethyl-benzyl-L-cysteine by the DCC procedure. Acetylation was achieved with acetic anhydride-pyridine for 1 h. All couplings were monitored by the ninhydrin reaction²⁶ and repeated until a negative test was observed. A portion of the protected peptide-resin (10.5 g) was cleaved with anhydrous hydrogen fluoride (containing 10% dithioethane) and the crude product (1.98 g) purified by preparative high pressure liquid chromatography (HPLC) on a Nucleosil C₁₈ reversed-phase column using a linear gradient of (A) H₂O (0.1% trifluoroacetic acid (TFA) to (B) acetonitrile (0.1% TFA) from 5%–25% (B) in 120 min. The purified peptide (914 mg, 19.6% yield) was homogenous by analytical HPLC. TT(350 mg) in 80 ml of 0.1 M phosphate buffer (pH 7.5) was reacted with *N*-succinimidyl-3-maleimidopropionate (586 mg) in dimethylformamide, adjusted to pH 6 (HCl) and applied to a Sephadex G25S column. When eluted with 0.05 M phosphate buffer (pH 6.0), the modified TT appeared as the first peak. This was reacted with Ac-Cys-(NANP)₃-OH (148 mg) and re-chromatographed on a Sephadex G25S column as above. The resultant conjugate was dialysed against saline and concentrated (50 ml).

tides conjugated to protein carriers generate specific anti-sporozoite antibodies in animals¹⁰. Accordingly, a vaccine suitable for testing in man was prepared by conjugating acetyl-cysteine-(NANP)₃-OH [Ac-Cys-(NANP)₃-OH] made by solid-phase peptide synthesis¹² to tetanus toxoid (see Table 1 for details). Amino acid analysis revealed 25 residues of peptide per mole of tetanus toxoid.

Male students from the University of Maryland Schools of Medicine and Dentistry were selected based on their expressed interest, good health, and lack of previous history of malaria. Medical screening included a history and physical examination, a complete blood count, urinalysis, and serology for hepatitis B, syphilis, and human immunodeficiency virus. Volunteers were excluded if they had received TT within the past 3 years, had known hypersensitivity to TT, or had previously experienced notable allergic reactions. All aspects of the studies were explained in detail, including potential risks and benefits, and informed consent was obtained. Studies were approved by ethical review committees of the University of Maryland Hospital and the National Institute of Allergy and Infectious Diseases.

Volunteers were randomized to four dosage groups to receive 20, 50, 100, or 160 µg of (NANP)₃-TT (Table 1). Vaccine was administered in double-blind fashion, with control volunteers within each group receiving a dosage of TT without peptide equal to the amount of TT in the conjugate vaccine for that dosage group. Booster vaccinations were administered 28, 56 and, in select volunteers, 133 days later.

Neither TT nor the conjugate vaccine elicited serious local or systemic untoward effects. The only common symptom was mild soreness at the vaccination site in 57% of vaccinees following the first injection; booster injections were similarly well tolerated. No specific reactions could be attributed to the synthetic peptide component of the conjugate vaccine.

The frequency and magnitude of the antibody responses correlated with the dose of vaccine administered (Table 2). Peak

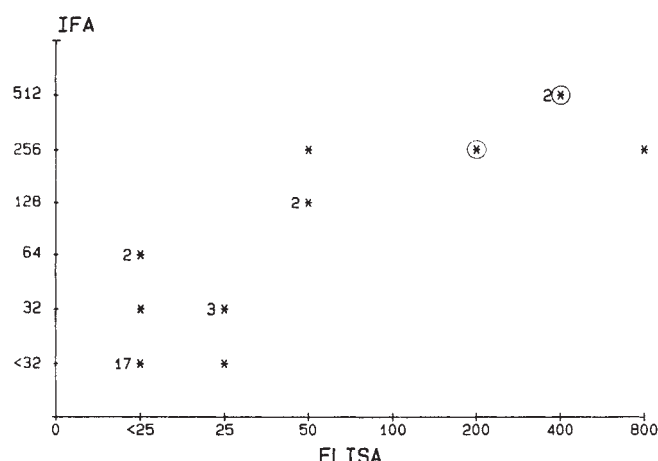


Fig. 1 Scatter plot of anti-peptide IgG (ELISA) against anti-sporozoite IgG (IFA) for recipients of all doses of (NANP)₃-tested on day +72. Volunteers participating in the sporozoite challenge study are circled. IFA was carried out using multiple well slides containing sporozoites from the NF54 strain of *P. falciparum* fixed in 0.1% glutaraldehyde²⁷.

ELISA titres to NANP were detected after the first dose of vaccine and did not rise further following booster immunizations. Most sera with antibody to NANP detected by ELISA reacted with *P. falciparum* sporozoites and the correlation between ELISA and IFA titres (Fig. 1) was significant ($r_s = 0.858$, $P < 0.001$). Circumsporozoite (CSP) reactions¹³, known to be poorly sensitive¹⁴, were not seen.

Peripheral blood lymphocytes from all vaccinees and controls were reacted with (NANP)₃ and (NANP)₅₀ in a lymphocyte replication assay¹⁵. Lymphocyte proliferation in response to NANP was not detected in vaccine recipients. The fact that the titre of antibody to peptide did not increase after booster vaccinations also indicates that T-cell stimulation by (NANP)₃ was not a major effect of immunization with (NANP)₃-TT.

Vaccine efficacy was assessed by inoculating volunteers with infectious sporozoites. *P. falciparum* strain NF54 was selected on the basis of known chloroquine sensitivity and high infectivity for mosquitoes, and was cultured as previously described¹⁶. *Anopheles stephensi* and *Anopheles freeborni* mosquitoes were infected by membrane feeding¹⁶ on a mixture of NF54 cultures and fresh normal blood with a final gametocytaemia of ~1–2%. Sporozoite inoculation was accomplished through the bites of infected mosquitoes. Each fully fed mosquito was immediately dissected to detect salivary gland sporozoites which were found in 50%. Salivary gland sporozoites were graded 1+ (1–10 per paired gland), 2+ (11–100 per paired gland), 3+ (101–1,000

per paired gland) or 4+ (>1,000 per paired gland)¹⁷. Each volunteer was exposed to five infected mosquitoes, at least one of which in each case had a 3+ or 4+ sporozoite density. This is a relatively high challenge dose since, except in areas of very high malaria endemicity, the infection rates of mosquitoes are generally much lower than 1%. Analysis of variance revealed no significant differences between individual volunteers in the mean sporozoite rates of the infected mosquitoes that bit them.

Giemsa-stained thick and thin blood films from each volunteer were examined every 12 hours, beginning 6 days after mosquito feeding. Parasitaemias were estimated by enumerating the number of parasites and leukocytes in a thick film until 500 leukocytes were counted. The number of parasites per microlitre of blood was determined based on the results of a white blood cell count made within 48 hours for each volunteer. A thick blood film was considered negative only after 15 or more minutes of examination, corresponding to at least 200 oil immersion fields. A course of oral chloroquine phosphate (600 mg base, followed by 300 mg 6, 24 and 48 hours later) was immediately initiated upon confirmation of parasitaemia.

Three vaccinees who had the highest rises in anti-sporozoite antibody by ELISA and IFA (Fig. 1) and four control volunteers with no history of prior malaria infection or vaccination participated in the efficacy study. As shown in Table 3, the four control volunteers developed parasitaemia 7, 8, 9 and 10 days following challenge (mean of 8.5 days). The two vaccinees who developed

Table 2 Summary of serological results (IgG and IgM-ELISA) of volunteers immunized with varying dosages of (NANP)₃-TT vaccine

Vaccine dose group	Number of vaccinees	Number (%) with significant rises in anti-(NANP) antibody			Day 0	Geometric mean IgG titre		
		IgG	IgM	IgG or IgM		Day 14	Day 44	Day 72
Controls	10	0	0	0	12.5	12.5	14.4	13.5
20 or 50 µg	6	0	0	0	12.5	12.5	12.5	14.9
100 µg	15	5 (33)	4 (27)	8* (53)	12.5	20.5	22.8	20.5
160 µg	14	8 (57)	3 (21)	10* (71)	13.1	35.4	41.0	42.6

Immunizations were given on days 0, 28 and 56 and sera were obtained at 2 week intervals following each vaccination. ELISA was carried out as follows: Wells of every other column of polystyrene Cooke ELISA plates (Dynatech Corporation) were coated with 100 ng per well of (NANP)₅₀ (synthetic peptide prepared in our laboratories) in carbonate coating buffer, pH 9.6. Plates were incubated overnight at 4 °C and washed 3 times with phosphate buffered saline (PBS) pH 7.3. PBS with 20% fetal bovine serum was added to each well. Plates were incubated for 1 h at 37 °C and washed three times with PBS containing 0.05% Tween. Serial twofold dilutions of sera were made starting at 1:25. Plates were incubated for 2 h at 37 °C, and washed 5 times with PBS containing 0.05% Tween. Goat anti-human IgG alkaline phosphatase conjugate (Kirkegaard Perry Laboratories) at a dilution of 1:600 in PBS containing 1% fetal bovine serum was added to each well. Plates were incubated for 1 h at 37 °C, washed 5 times as above and alkaline phosphatase substrate was added. Colour was permitted to develop for 30 min at 37 °C. The optical densities (OD) of the wells were read in a Titertek multiscan automated reader set at a wavelength of 450 nm. Net ODs were calculated by subtracting the OD of the corresponding background well from the antigen-containing test well. Titres were determined based on a cutoff OD corresponding to the mean OD of preimmunization sera at a 1:25 dilution plus 3 standard deviations. A four-fold rise in titre from pre- to post-immunization was considered significant. The procedure for determining IgM antibody was the same except that the conjugate used was goat anti-human IgM (Kirkegaard Perry Laboratories) at a dilution of 1:500.

* Sera from one volunteer contained both IgG and IgM antibodies to NANP.

Table 3 Serological and parasitological results after *P. falciparum* sporozoite inoculation in vaccinees and controls

Volunteer	Prevaccination	Reciprocal IgG anti-NANP titres		Day +28	Days to patency†	Peak parasitaemia (parasites mm ⁻³)
		Prechallenge	Day +7			
1	ND*	<25	<25	<25	9	49
2	ND	<25	<25	<25	8	20
3	ND	<25	<25	<25	10	<10
4	ND	<25	<25	<25	7	21
5	<25	400	200	400	11	24
6	<25	400	400	200	‡	‡
7	<25	200	400	400	11	24

Volunteers 1–4 were unimmunized controls. Volunteers 5, 6 and 7 were vaccine recipients. Antibody titres were measured by ELISA as previously described. Parasites mm⁻³ of blood were enumerated microscopically (see text).

* ND, not done. Control volunteers received no vaccine.

† $P=0.029$ when prepatent periods between vaccinees and controls are compared using 1-tailed Wilcoxon rank sum test.

‡ Volunteer 6 did not develop parasitaemia during the 29 day post challenge observation period.

malaria infection after challenge did not have asexual blood stage parasites detected until 11 days after challenge and the remaining vaccinee was still free of parasitaemia at 29 days; the difference in number of days until detection of parasitaemia between the controls and vaccinees is significant ($P=0.029$, Wilcoxon rank sum test). All infected volunteers received curative doses of chloroquine and recovered uneventfully; the maximum parasitaemia recorded was 49 parasites per mm³. The one non-infected volunteer received chloroquine after a 29-day period of observation. Red blood cells from this volunteer supported the growth of plasmodium strain NF54 in *in vitro* culture. Routine testing of urine¹⁷ showed that no 4-aminoquinolines had been surreptitiously ingested by any volunteer during the study.

Studies in sporozoite-induced malaria in man have suggested an inverse relationship between the size of the sporozoite inoculum and the duration of the prepatent period^{18,19}. This inverse relationship has been quantitatively documented in animal malaria^{20,21}. In mice infected with known numbers of dissected sporozoites of *P. berghei*, the prepatent period was lengthened when the inoculum was reduced from 5×10^3 to 10 sporozoites²². Thus, the significantly increased prepatent period of 11 days seen in the two vaccinees who developed infection compared to a mean of 8.5 days in the controls probably represents a significant vaccine-induced reduction in the numbers of sporozoites which entered and developed in hepatocytes.

These data demonstrate that a synthetic peptide of the immunodominant epitope of *P. falciparum* circumsporozoite protein conjugated to tetanus toxoid as a protein carrier has resulted in a first generation malaria vaccine that is safe in man, stimulates a biologically active immune response that apparently diminishes the number of circulating sporozoites that reach the liver and, in one of three vaccinees, provided complete protection against malaria. The serological response in mice immunized with similar vaccines incorporated in Freund's adjuvant and containing the repeated epitope of the *P. berghei* CS protein has been more prominent and has demonstrated the effect of booster immunizations. Moreover, most mice were protected against challenge although cell-mediated responses against the synthetic peptide could not be demonstrated²³. It is likely that improved formulations and new, more potent adjuvants, can markedly enhance the antibody response and perhaps also increase T-cell stimulation. It is also possible that tetanus toxoid, given to our volunteers in the past as childhood immunizations, dampened the immune response to the synthetic peptide component of the conjugate^{24,25}; further studies are planned to resolve this issue. We are also exploring whether HLA or other genetic restrictions to immune responsiveness to this vaccine exist. A possible disadvantage of this particular peptide conjugate vaccine is that it may not prime parasite-specific T cells, and a boosting of immune response and lymphokine production

may not occur during subsequent exposure to sporozoites under natural conditions². Nevertheless, the results reported here represent a step forward on the road towards the development of practical malaria vaccines.

After this paper was submitted, a report appeared²⁸ describing the results of a trial in humans of a recombinant DNA *P. falciparum* vaccine containing multiple NANP repeats. With this vaccine, protection also appeared to correlate with antibody levels.

We thank the US Agency for International Development, the National Institute of Allergy and Infectious Diseases, the American Institute of Biological Sciences, the MacArthur Foundation and Hoffman-La Roche for support. Absorption of (NANP)₃-TT to aluminium hydroxide and subsequent bottling of the vaccine were carried out at the Pasteur Institute under the supervision of Dr E. H. Relyveld. We thank Pamela Leland for technical assistance, and the volunteers for participating in the study.

Received 1 June; accepted 19 June 1987.

- WHO Technical Report Series, No 735, 1986 (WHO Expert Committee on Malaria: eighteenth report).
- Miller, L. H. *et al. Science* **234**, 1349–1356 (1986).
- Clyde, D. F., Most, H., McCarthy, V. C. & Vanderberg, J. P. *Am. J. med. Sci.* **266**, 169–177 (1973).
- Clyde, D. F., McCarthy, V. C., Miller, R. M. & Woodward, W. E. *Am. J. trop. Med. Hyg.* **24**, 397–401 (1975).
- Rieckmann, K. H., Beaudoin, R. L., Cassels, J. S. & Sell, K. W. *Bull. Wld Hlth Org.* **57**, (Suppl. 1), 261–265 (1979).
- Nussenzweig, V. & Nussenzweig, R. S. *Cell* **42**, 401–403 (1985).
- Zavala, F., Cochrane, A. H., Nardin, E. H., Nussenzweig, R. S. & Nussenzweig, V. *J. exp. Med.* **157**, 1947–1957 (1983).
- Dame, J. B. *et al. Science* **225**, 593–599 (1984).
- Enea, V. *et al. Science* **225**, 628–630 (1984).
- Zavala, F. *et al. Science* **228**, 1436–1440 (1985).
- Young, J. F. *et al. Science* **228**, 958–962 (1985).
- Barany, G. & Merrifield, R. B. *The Peptides: Analysis, Synthesis, Biology* Vol. 2 (Academic, New York, 1980).
- Vanderberg, J., Nussenzweig, R. S. & Most, H. *Mil. Med.* **134**, (Suppl.), 1183–1190 (1969).
- Nardin, E. H., Nussenzweig, R. S., McGregor, I. A. & Bryan, J. H. *Science* **206**, 597–599 (1979).
- Levine, M. M. *et al. J. clin. Invest.* **79**, 888–902 (1986).
- Burkot, T. R., Williams, J. L. & Schneider, I. *Trans. R. Soc. trop. Med. Hyg.* **78**, 339–341 (1984).
- Lelijveld, J. & Kortman, H. *Bull. Wld Hlth Org.* **42**, 477–479 (1970).
- Jeffery, G. M., Young, M. D., Burgess, R. W. & Eyles, D. E. *Annals trop. Med. Parasit.* **53**, 51–58 (1959).
- Boyd, M. F. *Symposium on Human Malaria* Pub. 15 (American Association for the Advancement of Science, Washington DC, 1941).
- Nardin, E. H. *et al. J. exp. Med.* **156**, 20–30 (1982).
- Schmidt, L. H., Fradkin, R., Genther, C. S., Rossan, R. N., Squires, W. & Hughes, H. B. *Am. J. trop. Med. Hyg.* **31**, (Suppl.), 609–703 (1982).
- Nussenzweig, R. S. *Expl Parasit.* **21**, 224–231 (1967).
- Zavala, F. *et al. J. exp. Med.* (submitted).
- Herzenberg, L. A., Tokuhis, T. & Herzenberg, L. *Nature* **285**, 665–667 (1980).
- Schutz, M. P., Leclerc, C., Audibert, F. & Chedid, L. *J. Immun.* **135**, 2319–2322 (1985).
- Sarin, V. K., Kent, S. B. H., Tam, J. P. & Merrifield, R. B. *Analyt. Biochem.* **117**, 147–157 (1981).
- Nardin, E. H., Gwadz, R. W. & Nussenzweig, R. S. *Bull. Wld Hlth Org.* **57** (Suppl. 1), 211–217 (1979).
- Ballou, R. S. *et al. Lancet* **i**, 1277–1281 (1987).

Evidence for a physical association of CD4 and the CD3: α : β T-cell receptor

Kaj Saizawa, Jose Rojo & Charles A. Janeway Jr

Section of Immunobiology and Department of Pathology, Howard Hughes Medical Institute at Yale University School of Medicine, New Haven, Connecticut 06510, USA

CD4 is a molecule expressed on the surface of T lymphocytes which recognize foreign protein antigens in the context of class II major histocompatibility complex (MHC) molecules^{1,2}. Recognition of antigen: class II MHC complexes by CD4⁺ T cells can be inhibited by anti-CD4 (ref. 3). Nevertheless, specific recognition of the antigen:Ia complex is clearly a function of the T-cell receptor, which is composed of CD3 and the variable polypeptides α and β ^{4,5}. Thus, it has been proposed that CD4 serves an accessory function in the interaction of CD4⁺ T cells and Ia-bearing antigen-presenting cells by binding to non-polymorphic portions of class II MHC molecules and stabilizing the cell interaction. Based on our observation that anti-CD4 could inhibit activation of a cloned line of CD4⁺ T cells by antibodies directed at a particular epitope on the variable region of the T-cell receptor, we have recently proposed that CD4 is actually part of the T-cell antigen recognition complex, physically associated with CD3: α : β ⁶. But numerous studies showing that CD3 and CD4 are not stably associated on the T-cell surface would appear to contradict this model. Here we show that anti-T-cell-receptor antibodies can co-modulate expression of the T-cell receptor and CD4, and that the monovalent Fab fragment of such an anti-T-cell-receptor antibody can, in conjunction with bivalent anti-CD4 antibody, generate an activating signal for the T cell. These findings provide further evidence for a physical association of the T-cell receptor complex and CD4.

For the experiments reported here we used a cloned CD4⁺, CD8⁻ AKR/J T-cell line called D10.G4.1 (D10) specific for the antigen conalbumin (CA) associated with the syngeneic I-A^k molecule⁷⁻¹². This cloned T-cell line will respond to a battery of monoclonal anti-T-cell-receptor antibodies⁷⁻⁹ in soluble form. Although the bivalent form of anti-receptor antibody is stimulatory for D10 cells, the monovalent Fab fragment is inhibitory^{8,9,12} of all specific stimuli, such as other monoclonal anti-receptor antibodies or CA:I-A^k. The activation of D10 cells by CA:I-A^k is readily inhibited by anti-CD4, whereas activation by most anti-T-cell-receptor antibodies is not^{6,10,11}. But responses of D10 cells to one set of antibodies (called 5A and 13A) directed at a distinct epitope on the variable region of the T-cell receptor are readily inhibited by anti-CD4 antibody⁶. This finding suggested that CD4 might be closely associated with the T-cell receptor, as it is otherwise difficult to understand why the epitope recognized by a particular anti-V region antibody should determine the biological effect of anti-CD4.

The interaction of 5A and a prototypic anti-D10 receptor antibody 3D3 with D10 cells is examined in Fig. 1. Both 3D3 and 5A bind well to D10 cells, as shown in Fig. 1a, and compete effectively with one another for binding to the receptor (Fig. 1b,c). As neither antibody binds to other cloned lines or to normal AKR/J T cells, both are directed at the T-cell receptor variable region; 3D3 is known to precipitate the α : β heterodimer of D10 cells¹². These antibodies are directed at physically distinct epitopes, however, as anti-CD3_e antibodies inhibit binding of 5A and 13A, but not of 3D3 (J.R. and C.A.J., manuscript in preparation). Anti-CD4 does not affect binding of any anti-T-cell-receptor antibodies (Fig. 1d and J.R. and C.A.J. unpublished data). When the biological activity of antibodies 5A and 3D3 is compared in terms of the concentration of antibody required to induce maximal IL-4 secretion by D10 cells, it can be seen (Fig. 1e) that 3D3 is very much more potent than 5A, even adjusting for the slight difference in binding

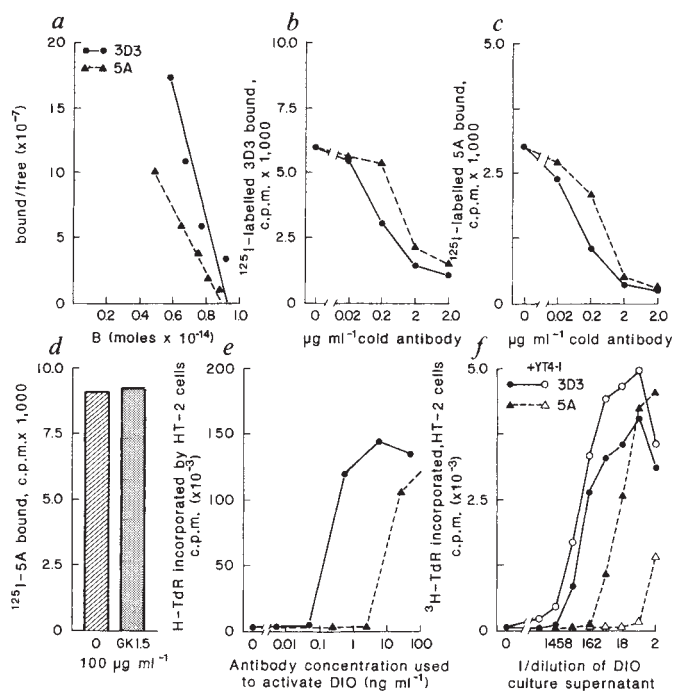


Fig. 1 Comparison of binding to and ability to activate cloned line D10.G4.1 by monoclonal anti-T-cell-receptor variable region antibodies 3D3 (○) and 5A(Δ). *a*, Scatchard binding curves of 125I-labelled monoclonal antibody to live D10.G4.1. *b*, *c*, Inhibition of binding of 1 μg ml⁻¹ 125I-labelled 5A or 3D3 by unlabelled 5A or 3D3. *d*, No effect of anti-CD4 mAb GK1.5 at 100 μg ml⁻¹ on binding 1 μg ml⁻¹ 125I-labelled 5A. *e*, *f*, Stimulation of 3H-thymidine (3H-TdR) incorporation by HT-2 indicator cells by supernatants of D10 cells stimulated with different concentrations of 5A or 3D3 for 24 h (*e*) or of 20 μg ml⁻¹ 3D3 or 5A with (○, Δ) or without (●, ▲) the anti-CD4 mAb YT4-1 at 5 μg ml⁻¹ (*f*), with supernatants taken at 8 h and titrated on HT-2 cells.

Methods. 5A and 3D3-containing ascites were purified by protein A-sepharose affinity column fractionation, labelled with 125I by the chloramine T method, and used for binding at various concentrations to 3 × 10⁵ D10 cells in 20 μl. Reactions were carried out on ice in 0.1% Na₂S₂O₃ for 2 h, and cells were washed four times before counting duplicate samples. Non-specific binding in the presence of a 100-fold excess of unlabelled homologous antibody was subtracted. Calculated *K_d* for 5A was 3.7 × 10⁻⁹ M⁻¹ and for 3D3 was 9.2 × 10⁻¹⁰ M⁻¹. *b*, *c*, *d*, Aliquots of 3 × 10⁵ D10 cells were incubated for 2 h with varying amounts of unlabelled monoclonal antibody (mAb) then 1 μg ml⁻¹ final 125I-labelled mAb was added. Binding was assayed as for *a*. *e*, D10 cells at 2 × 10⁴ per 200 μl culture were incubated 24 h with mAb, then 25 μl supernatant was added to 75 μl with 10⁴ HT-2 cells. Incorporation of 3H-TdR into triplicate cultures was measured by pulsing from 24 to 36 h after culture initiation. *f*, D10 cells at 10⁶ per ml were incubated for 8 h with 20 μg ml⁻¹ anti-T-cell receptor ± 5 μg ml⁻¹ YT4-1, and supernatants were titrated and added to HT-2 cells as for *e*. D10 cells were prepared as described in ref. 7.

affinity of the two antibodies. Finally, as had previously been shown for D10 proliferation, anti-CD4 slightly potentiates the response to 3D3, but actually inhibits the response of D10 cells to 5A by 20-fold in an assay carried out under the exact conditions used for the binding assays (Fig. 1f). These data show that both 5A and 3D3 can potently activate D10 cells when used at sufficient concentration, but that they differ significantly in their ability to do so as well as in the effect of anti-CD4 upon this response.

Recent studies by Emrich and colleagues¹³ have explored the role of CD8 in T-cell activation. They report that CD8⁺ T cells that do not respond to anti-CD3 will respond when anti-CD3 and anti-CD8 are physically linked to one another. They interpret their data to mean that optimal signalling of CD8⁺ T cells requires a physical association of these two

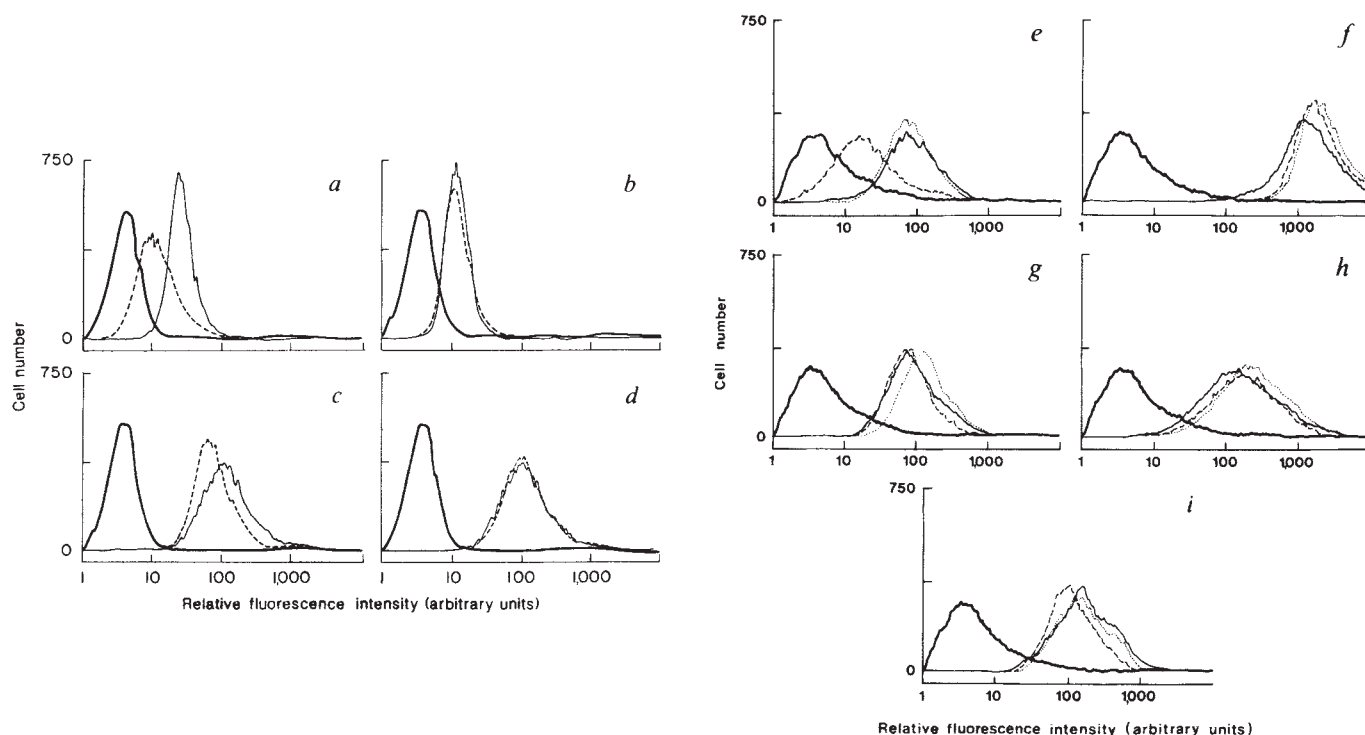


Fig. 2 FACS analysis of D10.G4.1 cells stained with 3D3-biotin (*a*), 5A-biotin (*b*), GK1.5-biotin (*c*, *d*) after overnight incubation of 5×10^5 D10 cells in 200 μ l containing 50 μ g ml $^{-1}$ 3D3-biotin (*a*), 5A-biotin (*b*), 3D3 (*c*) or 5A (*d*). All cells washed, stained with biotin-conjugated antibody as stated, and developed with FITC-conjugated Avidin (Vector Labs). Thick line, cells stained with FITC-avidin alone; thin line, cells incubated overnight without antibody and stained with biotin-mAb as stated; dashed line, cells incubated overnight with mAb or biotin-mAb, and restained with biotin-mAb. Panels *e*–*i*, experiment 2: D10 cells were incubated for 14 h at 37 °C in medium (thin line), 50 μ g ml $^{-1}$ 3D3 (dashed line), or 50 μ g ml $^{-1}$ 5A (dotted line), washed, stained with biotin antibodies and analysed in the FACS. Thick line, staining with FITC-avidin alone. *e*, Staining with 3D3 followed by biotin-rabbit anti-mouse immunoglobulin followed by FITC-avidin. *f*–*i*, Staining with biotin-monoclonal antibody followed by FITC-avidin. *f*, Y-3, anti-H-2K d ; *g*, M17, anti-LFA-1; *h*, anti-Tap, kindly provided by Dr Kenneth Rock 16 ; *i*, GK1.5, anti-CD4. 3D3 clearly modulates the T-cell receptor, and again co-modulates CD4, whereas 5A does neither. Neither anti-receptor antibody modulates H-2K d , LFA-1 or Tap. Thus, the modulation of CD4 is specific.

molecules, and that this is normally accomplished by having both molecules bind to the same class I MHC molecule on the stimulating cell.

In view of these findings and the differences between the different anti-T-cell-receptor antibodies shown in Fig. 1, we asked whether the high potency of 3D3 for D10 activation might reflect an induced association between the T-cell receptor and CD4 upon binding of 3D3 antibody. To examine this question, we took advantage of our earlier observation that 3D3 could modulate the receptor on D10 cells when used in soluble form 11 . As shown in Fig. 2*a*, overnight incubation of D10 cells with 3D3 does reduce 3D3 expression by ~50%. Staining these cells with anti-CD4 shows a significant reduction in CD4 expression as well, as seen in Fig. 2*c*. In a control experiment, shown in Fig. 2*e*–*i*, this result is repeated, and the expression of H-2K d , H-2D k , LFA-1 and Tap (a member of the Ly6 family of antigens) is seen not to change under conditions in which CD4 expression is reduced by incubation with 3D3. Thus, 3D3 causes a reduced expression of the T-cell receptor and of CD4 on D10 cells. This reduction in CD4 expression could be a consequence of D10 activation. But preincubation of D10 cells with high concentrations of 5A, sufficient to achieve maximal stimulation of D10 cells, does not alter expression of CD4; interestingly, 5A also fails to induce receptor modulation, for reasons that are presently not understood. This clearly dissociates activation from modulation of either the T-cell receptor or CD4, however. In a final control experiment, we have shown that antibodies to several other cell surface molecules, including CD4 itself, do not alter expression of either CD4 or the T-cell receptor, and incubation with PMA also fails to alter CD4 expression under these conditions. Thus, the co-modulation of CD4 with the

antigen receptor induced by 3D3 appears to be specific, and is thus likely to reflect an induced physical association between these two molecules.

The findings in Fig. 2 have been repeated on several occasions, giving the results shown in Table 1. Here, the percent reduction in antibody binding to the cell surface is shown as a function of the antibody used for preincubation. Both by FACS analysis as shown in Fig. 2, and by binding assay as shown in Fig. 1 (and J.R. and C.A.J., unpublished data), we know that D10 cells express 0.36 ± 0.1 3D3 binding sites per anti-CD4 binding site. Using the average data in Table 1 and this ratio, we can calculate that for each 3D3 binding site lost from the cell surface we lose about 2 anti-CD4 binding sites ($(1 \times 0.3)/(0.36 \times 0.5) = 1.7$). As these assays are carried out at saturating levels of antibody, if we assume that each molecule has a single binding site for each monoclonal antibody, then about two CD4 molecules are lost from the cell surface per T-cell receptor complex modulated. This suggests that CD4 may associate with the T-cell receptor as a dimer under these experimental conditions.

In further studies (not shown), we have examined about eight anti-D10 T-cell-receptor antibodies for their ability to co-modulate CD4. We find that those antibodies that are most potent for D10 activation, causing maximal activation under conditions in which few molecules of antibody are bound to the receptor, all cause co-modulation of the receptor and CD4, whereas those antibodies requiring high numbers of molecules bound to cause maximal activation do not co-modulate CD4 and the receptor. This latter group includes 13A, but also some antibodies directed at other epitopes, the response to which is not inhibited by anti-CD4. Thus, it would appear that high potency antibodies cause CD4 and the receptor physically to associate and to

Table 1 Modulation of CD4 molecules during T-cell receptor modulation

Pretreatment	Per cent decrease of T-cell receptor staining					s.d.
	Expt 1	Expt 2	Expt 3	Expt 4	Mean	
3D3-biotin	45.2	48.4	73.5	23.5	47.7	20.0
5A-biotin	13.5	0	0	4.1	4.4	6.4
GK1.5	0	0	0	0	0	0
	Per cent decrease of CD4 staining					s.d.
	Expt 1	Expt 2	Expt 3	Expt 4	Mean	
3D3	31.8	25.1	33.6	27.4	29.5	3.9
5A	4.4	9.2	0	4.3	4.5	3.8

D10 cells were purified by Ficoll-Hypaque centrifugation at least 10 days after stimulation. Aliquots of 5×10^5 cells were incubated with $50 \mu\text{g ml}^{-1}$ of indicated monoclonal antibody in Click's EHAA medium with 10% fetal calf serum overnight, washed, and stained with biotin-3D3 for T-cell receptors or biotin-GK1.5 under conditions giving plateau staining, washed, and stained with FITC-avidin, $100 \mu\text{g ml}^{-1}$. Cells were analysed on a FACS IV equipped with a Consort 40 data-processing unit. Relative fluorescence intensity was determined and compared to controls incubated overnight with medium alone. Percentage decrease in staining was calculated as:

$$100 - \frac{\text{Relative fluorescence intensity of test}}{\text{Relative fluorescence intensity of control}} \times 100$$

Mean and standard deviation of four experiments presented.

co-modulate. This result is consistent with the findings of Emmrich *et al.*¹³ cited above and recent findings of Fazekas de St. Groth (personal communication) in which CD4 or CD8 and the T-cell receptor are deliberately co-aggregated with bifunctional antibody, and suggest that the co-aggregation of CD4 with the T-cell receptor contributes to T-cell activation, requiring the cross-linking of far fewer receptor complexes to achieve optimal T-cell activation. In the absence of this co-aggregation of CD4 with the T-cell receptor, a higher degree of receptor cross-linking is required to achieve similar levels of T-cell activation.

If certain anti-T-cell-receptor antibodies act upon the T-cell receptor in such a way as to induce a physical association of the receptor with CD4 whereas others do not, then it seems possible that this could reflect a conformational change induced in the receptor as a result of antibody binding. To examine this hypothesis, we added the monovalent Fab fragment of 3D3 to D10 cells. If binding of this fragment induces an association of the receptor with CD4 molecules, then anti-CD4 might serve to provide the essential cross-linking function required for T cell activation, as CD4 would now be part of the receptor complex. As seen in Table 2, this combination of antibodies does indeed lead to D10 activation. Anti-LFA-1 and anti-class II MHC antibodies that react well with D10 cells do not induce a similar activation event. The 3D3-Fab fragment is, as previously reported, actually inhibitory for responses to intact 3D3, and anti-CD4 by itself has no stimulatory effect on D10 cells. Thus, this experiment further supports our hypothesis that certain high potency anti-T-cell-receptor antibodies induce a physical association of CD4 with the T-cell receptor, and that this complex of CD4 and the receptor can, upon cross-linking, generate a potent stimulatory signal for T-cell activation.

One initially puzzling aspect of these studies is that anti-CD4 in bivalent (Fig. 1f) or monovalent (S. Haque, K.S., J.R. and C.A.J., manuscript in preparation) will inhibit activation of D10 by 5A, an antibody not inducing a physical association with CD4, but does not inhibit activation by 3D3, an antibody that does induce this association. We believe that this apparent paradox can be readily understood in terms of the findings

Table 2 Synergy between bivalent anti-CD4 and nonvalent anti-T-cell-receptor antibody in T-cell activation

Bivalent antibody	Specificity	Concentration	Proliferative response of D10 cells in presence of	
			Medium	Monovalent 3D3-Fab
Experiment 1				
0	-	-	297	3,475
3D3	$\alpha : \beta$	4 ng ml ⁻¹	43,986	15,765
GK1.5	CD4	20 μ g ml ⁻¹	205	18,113
GK1.5-sepharose	CD4	-	506	14,845
Experiment 2				
0	-	-	3,460	9,332
3D3	$\alpha : \beta$	4ng ml ⁻¹	48,713	29,742
YT4-1	CD4	10 μ g ml ⁻¹	3,065	24,064
Y-3	H-2K ^k , D ^k	10 μ g ml ⁻¹	3,360	11,050
M17	LFA-1	10 μ g ml ⁻¹	3,058	10,582

D10 cells were cultured at 2×10^4 cells in 0.1 ml Click's EHAA with 10% fetal calf serum and 3% P388.D1-derived supernatant containing IL-1¹⁰. Added to the cells was the monovalent Fab fragment of 3D3 to achieve 200 ng ml^{-1} and the bivalent antibodies at the concentration shown. GK1.5 was coupled to sepharose beads at 3 mg protein per ml packed beads and $0.5 \mu\text{l}$ of packed beads were added to some cultures. The anti-H2K^b monoclonal antibody Y-3, specific for H-2K^b, D^k on D10 cells and the M17 antibody specific for LFA-1 were shown to bind strongly to D10 by FACS analysis and, in the case of M17, by inhibition of the response of D10 to antigen and H-2^k feeder cells.

reported above. First, we would argue that even activation by 5A involves cross-linking of the rare pre-existing complexes of the T-cell receptor and CD4, the cross-linking of T-cell receptor complexes alone having little biological effect. We would further propose that anti-CD4 disrupts such complexes when 5A is bound to them, either by steric inhibition or by conformational changes in CD4 induced by anti-CD4 that are incompatible with steric changes in the receptor generated by 5A. By contrast, 3D3 induces conformational changes in the receptor that favours association with CD4, and anti-CD4 serves to cross-link the CD4: T-cell receptor complexes further, having a slight augmenting effect under conditions of optimal activation, or a potent activating effect where cross-linking by 3D3 is prevented (Table 2). It should be noted that some antibodies that fail to co-modulate CD4 and which have a low potency for D10 activation are not inhibited by anti-CD4; these antibodies recognize epitopes distinct from that recognized by 5A and 13A. For this reason, we favour the hypothesis that anti-CD4 and 5A binding leads to a steric effect preventing CD4 and the T-cell receptor from generating the optimal stimulating complex.

Anti-T-cell-receptor antibodies are not the physiological activators of CD4⁺ T cells; antigen presenting cells displaying antigen: class II MHC complexes on the cell surface are. In light of the present findings, it seems likely that when the T-cell receptor binds to antigen: class II MHC on the cell surface, that CD4 molecules bind to the non-polymorphic portions of the same class II MHC molecule, thus forming a complex with the T-cell receptor. If, as suggested by the studies with D10 cells, this complex is about 30- to 100-fold more effective for T-cell activation than is cross-linking of the T-cell receptor complex alone, then it is not surprising that CD4 expression is so tightly associated with the recognition of antigen in association with class II MHC molecules. CD8 cannot subserve this function in class II MHC recognition, as shown clearly by recent studies of Fazekas de St. Groth *et al.*¹⁴ and by ourselves¹⁵ on class II MHC recognizing T cells expressing both CD4 and CD8. In our studies, we showed that anti-CD4 was 25- to 100-fold more potent at inhibiting class II MHC recognition by such double-positive T cells as was anti-CD8. This number is comparable to

the differences in potency of anti-T-cell-receptor antibodies that do or do not co-modulate CD4.

It has been observed that CD4⁺ T cells can respond to allogeneic or autologous class I MHC molecules. We have previously argued that alloreactivity involves recognition of unmodified MHC molecules expressed at high multiplicity on the antigen presenting cell surface¹². Thus, recognition of such unmodified class I MHC molecules would resemble activation of D10 by 5A, and would not require the colocalization of CD4. Indeed, our model predicts that it is precisely in such situations that CD4⁺ T cells should recognize class I MHC molecules.

In conclusion, the present studies provide three different pieces of data supporting the hypotheses that CD4 is a physical part of the T-cell receptor for antigen: class II MHC. First, antibodies directed at different V-region epitopes are differentially affected in their ability to activate a cloned T-cell line by anti-CD4. Second, certain anti-T-cell-receptor antibodies, which share the property of activating the cloned T-cell line at low numbers of molecules bound, cause a co-modulation of CD4 molecules with the T-cell receptor. It appears that two CD4 molecules are modulated per receptor molecule. Antibodies that do not co-modulate the receptor and CD4 will only activate the cloned T-cell line when 30 to 100 times more antibody is bound per cell. Third, the monovalent Fab fragment of a high potency anti-receptor antibody, although itself unable to activate the cloned T-cell line, does render the cell activatable by means of bivalent anti-CD4, suggesting that CD4 is induced to become

a part of the receptor complex by conformational changes induced in the receptor by the anti-receptor Fab fragment. We propose that optimal T-cell activation involves CD4 and the T-cell receptor binding to the same antigen: class II MHC complex on the antigen presenting cell surface, and that the 30- to 100-fold increase in signalling observed when CD4 is part of the T-cell receptor complex can account for the tight association of CD4 expression and recognition of antigen in the context of self class II MHC molecules.

Received 5 March; accepted 13 May 1987.

1. Schwartz, R. H. *A. Rev. Immun.* **3**, 237-261 (1985).
2. Strominger, J. L. *Prog. Immun.* **4**, 541-554 (1980).
3. Dialynas, D. P. *et al. Immun. Rev.* **74**, 29-56 (1983).
4. Marrack, P. & Kappler, J. *Adv. Immun.* **38**, 1-30 (1986).
5. Dembic, Z. *et al. Nature* **320**, 232-238 (1986).
6. Janeway, C. A., Jr., Haque, S., Smith, L. A. & Saizawa, K. M. *J. molec. Cell. Immun.* **3**, 121-131 (1987).
7. Kaye, J., Porcelli, S., Tite, J., Jones, B. & Janeway, C. A. *Jr J. exp. Med.* **158**, 836-856 (1983).
8. Tite, J. P. *et al. J. exp. Med.* **163**, 189-202 (1986).
9. Janeway, C. A. Jr, Broughton, B., Smith, L. A., Marion, T. N. & Bottomly, K., *J. molec. Cell. Immun.* **3**, 83-94 (1987).
10. Kaye, J. *et al. J. Immun.* **133**, 1339-1345 (1984).
11. Kaye, J., Jones, B. & Janeway, C. A. *Immun. Rev.* **81**, 39-63 (1984).
12. Kaye, J. & Janeway, C. A. *Jr J. exp. Med.* **159**, 1397-1412 (1984).
13. Emrich, F., Strittmatter, U. & Eichmann, K. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8298-8302 (1986).
14. Fazekas de St. Groth, B., Gallagher, P. F. & Miller, J. F. A. P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2594-2598 (1986).
15. Jones, B., Khavari, C. C. & Janeway, C. A., Jr *J. Immun.* (in the press).
16. Rock, K. L. *et al. J. exp. Med.* **163**, 313-333 (1986).

Regulation of T-cell receptor γ -chain RNA expression in murine Thy-1⁺ dendritic epidermal cells

William A. Kuziel*, Akira Takashima†, Mark Bonyhadil‡, Paul R. Bergstresser†‡, James P. Allison||, Robert E. Tigelaar†‡ & Philip W. Tucker§

*Graduate Program in Immunology and Departments of

†Dermatology, ‡Internal Medicine and §Microbiology, University of Texas Health Science Center at Dallas, Dallas, Texas 75235, USA

|| Cancer Research Laboratory and Department of Microbiology and Immunology, University of California, Berkeley 94720, USA

The epidermis of normal mice contains two distinct populations of dendritic cells derived from the bone marrow, Ia⁺ Langerhans cells and Ia⁻ cells that express the Thy-1 alloantigen¹⁻⁵. The Thy-1-bearing dendritic epidermal cells (Thy-1⁺ dEC) have a surface phenotype similar to that of very early T-lineage cells⁶⁻⁸, produce IL-2-like growth factors¹⁰ and exhibit cytotoxicity which is not restricted by the major histocompatibility complex (MHC)^{10,11}. The relationship of Thy-1⁺ dEC to the T-cell lineage is unclear. Most T lymphocytes bear a receptor for antigen composed of an α chain and a β chain¹²⁻¹⁷ associated with a nonpolymorphic complex termed CD3 (T3)¹⁸⁻²¹. A minor population carries a receptor in which CD3 is associated with a γ/δ complex²¹⁻²⁵. We have analysed clones of Thy-1⁺ dEC for rearrangement and expression of the genes for the α -, β - and γ -chains of the T-cell receptor (TCR). They do not express α or β but do carry a γ/δ complex. Activation of the cells with Con A is associated with a rapid decrease in the steady-state level of γ -chain RNA. Because Thy-1⁺ dEC resemble early stage T lymphocytes, down-regulation of TCR expression may reflect a necessary event during T cell differentiation.

In Thy-1.1 mouse strains, such as AKR/J Thy-1⁺ dEC and the vast majority of thymocytes express the Thy-1-related determinant 20-10-5S, which is not readily detectable on peripheral

T cells^{9,26}. Several Thy-1⁺ dEC clones were established from AKR/J epidermal cell suspensions by limiting-dilution microculture of flow-cytometry-purified 20-10-5S⁺ cells in the presence of Con A, recombinant human IL-2 and X-irradiated syngeneic filler cells. Clones were expanded and maintained by feeding IL-2 every 3-4 days and restimulating with Con A at 3-8 week intervals. Three such clones (2E1, 2H3 and 1D2), and one line (termed 7-17) of Thy-1⁺ dEC established by stimulating bulk cultures of AKR/J epidermal cells with Con A and IL-2, were selected for further study. All populations were >98% Thy-1⁺ and <1% L3T4⁺; 2E1, 2H3 and 7-17 were <1% Lyt-2⁺ while 1D2 contained 2-3% dimly staining Lyt-2⁺ cells¹⁰.

Northern blot analysis (Fig. 1a) showed that none of these populations expressed detectable TCR α -chain RNA. Only 2E1 produced full-size, 1.3-kilobase (kb) β -chain RNA but in very low amount. All four populations produced truncated, 1.0-kb β -chain RNA and 1D2 also had a prominent 1.15-kb species of β -chain RNA. In contrast, all four populations expressed abundant levels of potentially productive (1.5-1.7 kb) TCR γ -chain RNA when probed with the constant region of C γ 1 (see Fig. 1b). Analysis with a V_{1.2} variable region probe revealed that each population expressed at least one of the functional V₁-family-associated C γ genes, either V_{1.1}-C γ 4 or V_{1.2}-C γ 2. The slight shift upward of the C γ band in 2E1 after hybridization with the V_{1.2} probe suggested that this clone expresses abundant V_{1.1}-C γ 4 RNA, as Iwamoto *et al.*²⁷ have shown that the C γ 4 constant region is 69 base pairs (bp) larger than C γ 2. Expression of C γ 4 RNA in 2E1 was confirmed by hybridization with a C γ 4 constant region DNA probe (data not shown). All four populations contained both 1.7-kb and 1.5-kb forms of C γ 1 RNA exclusively with the proximal variable region, V₃; no hybridization was observed using probes for the upstream V₄ and V₂ regions (data not shown). The difference between the 1.7- and 1.5-kb species resides in the 3'-untranslated region, because the 1.5-kb band disappears upon hybridization to a probe corresponding to the last 180 bp of the C γ 1 gene. This use of alternative polyadenylation sites about 250 bases apart in C γ 1 transcripts (Fig. 1b) was previously observed in cDNAs generated from fetal thymocyte RNA²⁸. Hybridization with the 3'-UT probe also revealed low levels of a 1.6-kb transcript in 1D2, 2H3 and 2E1, which is probably a truncated RNA as it was not detected

with any of the V-region probes. All four cell populations express RNA for the δ chain of CD3.

To determine whether the cells expressed CD3-associated proteins at the cell surface, lysates of radiolabeled cells from line 7-17 were subjected to immunoprecipitation with a monoclonal antibody to murine CD3 ϵ (ref. 29). As shown in Fig. 2a, upon SDS-PAGE under non-reducing conditions a single band was obtained slightly greater in size than that of the α/β antigen receptor heterodimer of C6VL lymphoma cells¹². But unlike the C6VL α/β heterodimer, reduction of the 7-17 complex yielded two components of relative molecular mass (M_r) 52,000 (52 K) and 42 K. To determine whether either of these components might be related to the γ -chain gene, immunoprecipitations were also carried out with a rabbit antiserum raised to a peptide corresponding to the seven carboxyl-terminal residues of the murine $C_{\gamma 1}$, $C_{\gamma 2}$ and $C_{\gamma 3}$ gene products³⁰. As shown in Fig. 2b, immunoprecipitates obtained with the anti- γ serum contained the same 52K and 42K components obtained with anti-CD3 ϵ . When the CD3-associated components were solubilized and reprecipitated with the anti- γ -chain serum, the primary band detected upon SDS-PAGE was the 42K band. These results strongly suggest that 7-17 cells express a CD3-associated heterodimer composed of a 42K γ chain linked via disulphide bonds to a 52K δ chain. Analysis of two Thy-1⁺ dEC clones, 1D2 and 6G3 (which have TCR RNA profiles similar to that of 7-17) gave similar results (data not shown). The sizes for both AKR/J γ and δ chains are larger than those for γ - and δ -proteins from C57BL/6 fetal and immature thymocytes (35K and 45K, respectively)^{21,22}. Whether these size differences result from differences in primary structure and/or post-translational modifications remains to be determined.

Southern blot analysis of 2H3 and 2E1 DNA revealed extensive rearrangement of TCR γ -chain and β -chain genes (Fig. 3). *Eco*RI digestion showed that both clones had rearranged both $C_{\gamma 1}$ alleles (16–18 kilobase pairs (kbp)). 2H3 rearranged both $C_{\gamma 2}$ alleles (10.5–17 kbp) and 2E1 rearranged one $C_{\gamma 2}$ allele (10.5–15 kbp) and at least one $C_{\gamma 4}$ allele (10.8–9.5 kbp), consistent with their RNA profiles. The rearranged $C_{\gamma 1}$ *Eco*RI band (18 kbp) contains the V_3 and V_4 regions, but the V_2 region is in the germline configuration (data not shown). Because rearrangements in the β -chain locus appear to correlate with cell maturation state, these genes were analysed in detail as described by Born *et al.*³¹. In these analyses, D_{β} to J_{β} partial rearrangements can be distinguished because they result in restriction fragments of predictable size. The data show deletion of both $C_{\beta 1}$ alleles from 2E1 and one $C_{\beta 1}$ allele from 2H3. These restriction patterns are consistent with each clone having at least one *DJ* partial rearrangement and one rearrangement which may represent complete *VDJ* joining. One possibility for the low level or lack of 1.3-kb β -chain RNA in 2E1 and 2H3, therefore, may be regulation of expression subsequent to rearrangement. The C_{α} constant region is present on a 12-kbp *Bgl*II restriction fragment in both clones. The state of rearrangement is not known, but the lack of TCR α -chain RNA in 2H3 and 2E1 cannot be attributed to deletion of both C_{α} alleles.

The surface phenotype and TCR RNA profile of Thy-1⁺ dEC closely resembles that of the CD3⁺4⁺8⁺, TCR (α/β)⁺ (γ/δ)⁺ cells described in human peripheral blood and in human and mouse thymus^{21–25}. There are other similarities between these cell populations. The human thymocyte clone CII secretes IL-2 and proliferates after appropriate stimulation with anti-CD3 (ref. 23). Con A stimulation of the Thy-1⁺ dEC line 7-17 induces the appearance of IL-2 RNA with a short-lived peak occurring between 6–9 h (Fig. 4b). Such stimulated cultures subsequently proliferate vigorously and culture supernatants support the growth of IL-2-dependent T-cell lines¹⁰. Several of the human cell populations display non-MHC-restricted cytotoxicity after culture with IL-2 or anti-CD3 antibody^{23,32} as do cultured Thy-1⁺ dEC lines and clones^{10,11}. IL-2 receptors were not detected on either freshly isolated CD3⁺, TCR(α/β)⁺ human cells³³

or on freshly isolated Thy-1⁺ dEC⁶, but after Con A stimulation of Thy-1⁺ dEC, cell surface IL-2 receptors are detected (data not shown). 7-17 Cells maintained for three weeks in media containing 10 U ml⁻¹ IL-2 express low but detectable levels of IL-2 receptor RNA which increase rapidly following restimula-

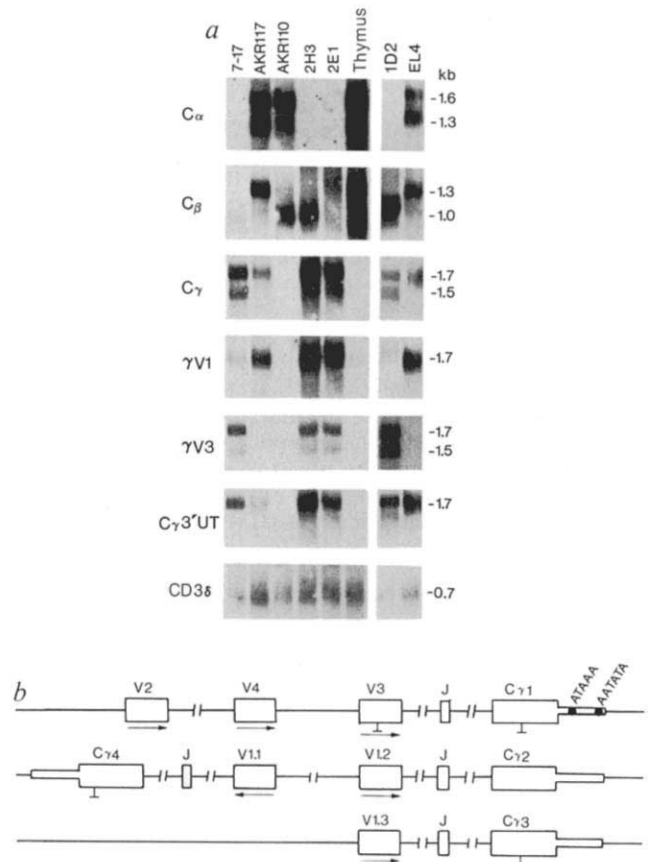


Fig. 1 a, T-cell receptor RNA profile of Thy-1⁺ dEC. The probes are listed at the left, the RNA sources are indicated above each lane, and RNA sizes (in kb) are indicated on the right. 7-17 is a Thy-1⁺ dEC line and 2H2, 2E1 and 1D2 are Thy-1⁺ dEC clones. AKR117, AKR110 and EL4 are thymoma tumour cell lines; thymus is 6 wk AKR thymus. b, Diagram and nomenclature of the murine TCR γ genes. The map was compiled from information in refs 27, 28, 30 and 47. $C_{\gamma 1}$ is associated with three non-cross-hybridizing variable (V) region genes; their relative positions as indicated are based on linkage of V_3 and V_4 and on deletion mapping results: V_4 and V_3 are deleted in cell lines that have rearranged both V_2 genes to $J_{\gamma 1}$ (ref. 31 and our unpublished observations). The two polyadenylation sites (ATAAA and AATATA) in $C_{\gamma 1}$ are 256 bp apart^{28,30}. $C_{\gamma 1}$ cross-hybridizes with $C_{\gamma 2}$ and $C_{\gamma 3}$ but not with $C_{\gamma 4}$. The cross-hybridizing V family $V_{1.1}$, $V_{1.2}$ and $V_{1.3}$ are associated with $C_{\gamma 4}$, $C_{\gamma 2}$ and $C_{\gamma 3}$, respectively. Because $C_{\gamma 3}$ has a potential splicing defect, it is considered to be a pseudogene. Horizontal arrows denote direction of transcription and inverted crosses in $C_{\gamma 1}$, $C_{\gamma 3}$, $C_{\gamma 4}$ and V_3 potential N-linked glycosylation sites. **Methods.** Total RNA from each cell population (~10 μ g) was fractionated by formaldehyde/agarose gel electrophoresis, transferred to Gene Screen (NEN) and hybridized to gel-purified DNA fragments ³²P-labelled as described previously⁴⁸. The same filter was used for each hybridization; the filter was submerged in 100 °C water for 5 min to remove the previously hybridized probe, as verified by autoradiography. The C_{α} and C_{β} probes were described previously⁴⁸. The C_{γ} probes were derived from a $C_{\gamma 1}$ cDNA called p γ 7.1 (our unpublished results); the 3'-untranslated (UT) probe is a *Pst*I-*Eco*RI fragment that corresponds to the 3'-terminal 180 bp of the $C_{\gamma 1}$ gene. The V region probes are described in ref. 30 and were provided by Dr D. Raulet. The CD3 δ probe is a 600-bp fragment encompassing almost the entire CD3 δ cDNA⁴⁹, and was provided by Dr C. Terhorst.

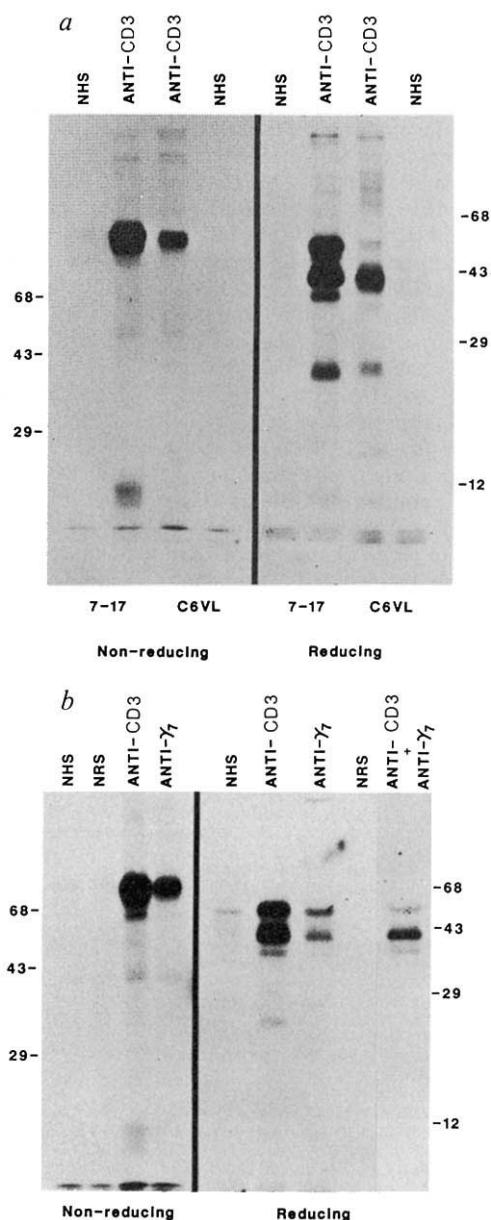


Fig. 2 Surface expression of CD3-associated γ/δ heterodimer by Thy-1⁺ dEC. 7-17 dEC or C6VL lymphoma cells were radioiodinated using lactoperoxidase as previously described¹² and solubilized in lysis buffer consisting of 5 mM CHAPS (3-[(cholamidopropyl)-dimethylammonio]-2,2-hydroxyl-1 propane-sulphonate) in 10 mM Tris-HCl, pH 8.0, and containing 1 mM phenylmethylsulphonyl fluoride, aprotinin and 10 mM iodoacetamide. Cell lysates were precleared with fixed protein A-bearing *Staphylococcus aureus* (SACI) coated with normal hamster serum (NHS) or normal rabbit serum (NRS). Immunoprecipitations were carried out with NHS or NRS, anti-CD3 monoclonal antibody (using culture supernatants from hybridoma 145-2C11 (ref. 29), or with a rabbit antiserum to a synthetic peptide corresponding to the seven carboxyl-terminal residues of the C_γ sequence. For sequential immunoprecipitation, CD3 immunoprecipitates were washed, and CD3-associated components eluted with lysis buffer containing 0.1% SDS and 1 mM dithiothreitol. After 5 min at 68 °C, iodoacetamide was added to a final concentration of 15 mM, the sample diluted with four volumes 1.5% NP-40 in 0.01 M Tris-HCl, pH 8.0, and the eluted proteins reprecipitated with anti- C_γ for 1 h. In all cases, samples were solubilized in SDS sample buffer with or without reducing agent, and subjected to electrophoresis on 10% (for non-reduced samples) or 12.5% (for reduced samples) SDS-PAGE gels. Gels were dried and labelled proteins visualized by autoradiography at -70 °C with enhancing screens.

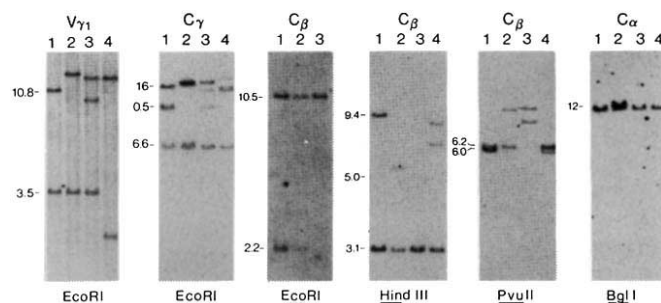


Fig. 3 Rearrangement of TCR γ and β genes in Thy-1⁺ dEC clones 2H3 and 2E1. Each panel shows a Southern blot of DNA digested with the indicated restriction enzyme and probed with the indicated TCR gene fragment (described in the legend to Fig. 1). Lane 1, AKR/J liver; lane 2, clone 2H3; lane 3, clone 2E1; lane 4, AKR thymoma cell line AKR110. Sizes of germline TCR gene restriction fragments (in kbp) are indicated at the left of each panel. In AKR/J DNA, these are: for γ , *EcoRI* 10.8 ($V_{\gamma 1}$, $V_{\gamma 2}$ and $C_{\gamma 4}$), *EcoRI* 3.5 ($V_{\gamma 3}$), *EcoRI* 16 ($C_{\gamma 1}$), *EcoRI* 10.5 ($C_{\gamma 2}$), *EcoRI* 6.6 ($C_{\gamma 3}$); for β , *EcoRI* 10.5 ($C_{\beta 2}$), *EcoRI* 2.2 ($C_{\beta 1}$), *HindIII* 9.4 ($C_{\beta 1}$), *HindIII* 3.1 ($C_{\beta 2}$), *PvuII* 6.2 ($C_{\beta 1}$), *PvuII* 6.0 ($C_{\beta 2}$); for C_α , *BglI* 12 (C_α).

Methods. DNA was prepared by the method of Blin and Stafford⁵⁰. DNA (10 μ g) was digested with 5-fold excess of restriction enzyme and fractionated through 0.8% agarose, transferred to nitrocellulose filters and hybridized for 24–36 h to ³²P-labelled DNA probes at 42 °C in 50% formamide, 5 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.2% polyvinylpyrrolidone, 0.02% Ficoll and 100 μ g ml⁻¹ denatured herring sperm DNA. Filters were washed in 3 $\times$ SSC, 0.01% SDS at room temperature, then in 0.1 $\times$ SSC, 0.1% SDS at 42 °C, air-dried and exposed to X-ray film at -70 °C using an intensifying screen.

tion with Con A and which remain elevated for 24 h. (Fig. 4).

The level of CD3 δ RNA also increased markedly within 12 h of stimulation with Con A (Fig. 4). But this treatment was also associated with a rapid and dramatic decrease in TCR γ -chain RNA in which both the V_3 - $C_{\gamma 1}$ and $V_{1.2}$ - $C_{\gamma 2}$ RNA species had decreased >4-fold by 12 h; prestimulation levels were not reattained by 24 h. The down regulation of C_γ RNA following mitogen activation of 7-17 contrasts sharply with the increases in RNA for other lymphoid-related genes such as CD3 δ , IL-2 and IL-2 receptor and cellular proteins such as β actin (Fig. 4). In other studies activation of human TCR (α/β)-bearing thymocytes was associated with an increase in β -chain RNA³⁴. The decrease in C_γ RNA is reminiscent of the decrease in RNA for immunoglobulin δ (IgD) heavy chain after stimulation of mature (IgM⁺, IgD⁺) B lymphocytes with lipopolysaccharide (LPS)³⁵ and the LPS-induced decrease in immunoglobulin V_H germline transcripts in Abelson murine leukaemia virus-transformed pre-B cell lines³⁶; in both cases RNA for IgM increases. These changes following mitogen activation *in vitro* reflect normal ontogenic differentiation *in vivo*. The decrease in IgD heavy-chain RNA occurs as a mature B cell differentiates into an IgM-secreting cell. Germline immunoglobulin V_H transcripts are most abundant in fetal liver and decrease to low levels in newborn or adult bone marrow³⁶. Similarly, T-cell γ -gene expression is substantially higher in fetal thymus than in newborn or adult thymus^{30,37–39}. Furthermore, steady-state levels of RNA containing the V_3 variable region of $C_{\gamma 1}$ are most abundant in 16–17-day-old fetal thymocytes and are not readily detectable in newborn or adult thymocytes, in resting or activated peripheral T cells or in other T-cell lines^{28,30}.

One explanation for the striking Con A-induced down-regulation of γ -chain RNA is that Thy-1⁺ dEC are an extrathymic equivalent of incompletely differentiated thymocytes, and Con A is mimicking an intrathymic differentiating signal. Con A has

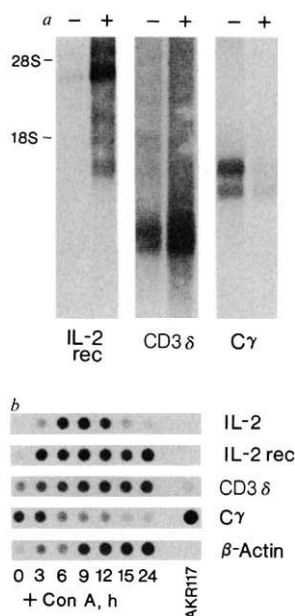


Fig. 4 Effect of Con A on RNA expression in Thy-1⁺ dEC line 7-17. *a*, Northern blot of RNA from untreated 7-17 cells (-) and from 7-17 cells treated with 1.5 $\mu\text{g ml}^{-1}$ of Con A for 12 h (+). Probes are indicated at the bottom of each panel; 28S and 18S ribosomal RNA markers are indicated on the left. *b*, Kinetics of RNA changes following Con A stimulation. The probes are listed at the right and time (in h) after Con A (1.5 $\mu\text{g ml}^{-1}$) addition is indicated at the bottom. The 7-17 cells used in both *a* and *b* had been last stimulated with Con A 21 days previously and maintained in 5-10 U ml^{-1} IL-2.

Methods. *a*, Methods for Northern blot analysis are described in Fig. 1. Total RNA (5 μg) was loaded in each lane. The same filter was used for all hybridizations after stripping in 100 °C water. *b*, Cytoplasmic extracts for the dot blot were prepared from 1.5×10^6 viable cells at each time point as described by White and Bancroft⁵¹; 50% of each extract was spotted through pores of a 96-well manifold apparatus onto nitrocellulose, and hybridized to ³²P-labelled probes as described for the Southern blots in Fig. 3. The same filter was used for each hybridization after stripping in 100 °C water. The mouse IL-2 probe is a 400 bp *HindIII*-*AccI* fragment from the 3'-end of the cDNA⁵² (provided by Dr F. Lee). The mouse IL-2 receptor probe is a 500-bp *PvuII*-*EcoRI* fragment corresponding to the 3'-end of the cDNA⁵³ (provided by Dr R. Germain). The chicken β -actin probe is a 700-bp *PvuII* fragment from the cDNA (provided by Dr D. Cleveland).

been shown to bind to the CD3-TCR (α/β) complex on peripheral T cells⁴⁰⁻⁴², and it is likely that this interaction delivers activation signals. It is possible that Con A is activating Thy-1⁺ dEC via the alternate CD3-TCR (γ/δ) complex. If cross-linking of such a complex is necessary for ultimate expression of the CD3-TCR (α/β) complex, the observed Con A-induced changes in RNA for CD3 δ and C γ might be expected; decreasing the amount of γ chains but increasing CD3 to accommodate the subsequent synthesis of α and β chains. One prediction of such a hypothesis is that these dendritic cells could be induced to undergo further maturation, as manifested, for example, by productive expression of TCR α - and β -chain genes, if the appropriate environment and/or stimulus is provided. On the other hand, the present findings are equally compatible with the possibility that Thy-1⁺ dEC are a distinct lineage of cells of unknown function whose response to activation signals includes transient down regulation of at least one of their TCR genes.

Another question raised by these studies is the relationship between Thy-1⁺ dEC and the phenotypically and genotypically similar cells found in the thymus (especially 15-17-day-old fetal

thymus). The characterization of the Thy-1⁺ dEC derived from the epidermis of nu/nu mice (Nixon-Fulton *et al.*, manuscript in preparation), indicates that these cells may not be post-thymic or thymus-dependent, but a pre-thymic population which has migrated to the epidermis. Similarities between the thymic and epidermal microenvironments have been noted previously⁴³⁻⁴⁶. The striking similarities between 15-17-day-old fetal thymocytes and Thy-1⁺ dEC raise the possibility that homologies between epidermis and fetal thymus may be particularly strong. Epidermally-derived populations of mouse Thy-1⁺ cells should facilitate analysis of the relationship of various populations of CD3⁺, TCR (γ/δ)⁺ cells to each other and to CD3⁺, TCR (α/β)⁺ cells, as well as the mechanisms involved in regulation of TCR γ gene expression.

We thank Drs D. Raulet, W. Born, R. Germain, F. Lee, D. Sherman, C. Terhorst and D. Cleveland for DNA probes, Dr P. Krammer for the AKR110 and AKR117 cell lines, Dr J. Bluestone for the gift of hybridoma 2C11 and Dr S. Horvath for the C γ 7 peptide. Ms Chhaya Das provided excellent technical assistance. This research was supported by the NIH.

Note added in proof: Since submission of this paper, König *et al.*⁵⁴ also have documented expression of TCR γ/δ proteins on Thy-1⁺ dEC.

Received 20 January; accepted 22 May 1987.

- Bergstresser, P. R., Tigelaar, R. E., Dees, J. H. & Streilein, W. J. *J. invest. Derm.* **81**, 286-288 (1983).
- Tschachler, E. *et al. J. invest. Derm.* **81**, 282-285 (1983).
- Breathnach, S. M. & Katz, S. I. *J. invest. Derm.* **83**, 74-77 (1984).
- Bergstresser, P. R., Tigelaar, R. E. & Streilein, J. W. *J. invest. Derm.* **83**, 83-87 (1984).
- Nixon-Fulton, J. L., Witte, P. L., Tigelaar, R. W., Kumar, V. & Bergstresser, P. R. *J. Immun.* **138**, 2902-2905 (1987).
- Romani, N. *et al. J. exp. Med.* **161**, 1368-1383 (1985).
- Bergstresser, P. R., Sullivan, S., Streilein, J. W. & Tigelaar, R. E. *J. invest. Derm.* **85**, 85s-90s (1985).
- Caughman, S. W. *et al. J. invest. Derm.* **86**, 615-624 (1986).
- Nixon-Fulton, J. L., Bergstresser, P. R. & Tigelaar, R. E. *J. Immun.* **136**, 2776-2786 (1986).
- Takashima, A., Nixon-Fulton, J. L., Bergstresser, P. R. & Tigelaar, R. E. (manuscript in preparation).
- Nixon-Fulton, J. L. *et al.* (manuscript in preparation).
- Allison, J. P., McIntyre, B. W. & Bloch, D. J. *J. Immun.* **129**, 2293-2300 (1982).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Acuto, O. *et al. Cell* **34**, 717-725 (1983).
- Kapler, J. *et al. Cell* **35**, 295-302 (1983).
- McIntyre, B. & Allison, J. P. *Cell* **38**, 659-665 (1984).
- Borst, J., Prendville, M. A. & Terhorst, C. *J. Immun.* **128**, 1560-1565 (1982).
- Samelson, L. E., Harford, J. B. & Klausner, R. D. *Cell* **43**, 223-231 (1985).
- Oettingen, H. C., Pettet, C. L., Meloy, W. L. & Terhorst, C. *Nature* **320**, 272-275 (1986).
- Lew, A. M. *et al. Science* **234**, 1401-1405 (1986).
- Pardoll, D. M. *et al. Nature* **326**, 79-81 (1987).
- Bank, I. *et al. Nature* **322**, 179-184 (1986).
- Brenner, M. B. *et al. Nature* **322**, 145-149 (1986).
- Lanier, L. L. *et al. J. exp. Med.* **165**, 1076-1094 (1987).
- Auchincloss, H., Ozato, K. & Sachs, D. H. *J. Immun.* **128**, 1584-1589 (1982).
- Iwamoto, A. *et al. J. exp. Med.* **163**, 1203-1212 (1986).
- Heilig, J. S. & Tonegawa, S. *Nature* **322**, 836-840 (1986).
- Leo, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 1374-1378 (1987).
- Garman, R. D., Doherty, P. J. & Raulet, D. H. *Cell* **45**, 733-742 (1986).
- Born, W., Yague, J., Palmer, E., Kappler, J. & Marrack, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2925-2929 (1985).
- Lanier, L. L., Ruitenberg, J. J. & Phillips, J. H. *J. exp. Med.* **164**, 339-344 (1986).
- Lanier, L. L. & Weiss, A. *Nature* **324**, 268-270 (1986).
- Ramarli, D., Fox, D. A., Milanese, C. & Reinherz, E. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7008-7012 (1986).
- Yuan, D. & Tucker, P. W. *J. Immun.* **132**, 1561-1565 (1984).
- Yankopoulos, G. D. & Alt, F. W. *Cell* **40**, 271-281 (1985).
- Raulet, D. H., Garman, R. D., Saito, H. & Tonegawa, S. *Nature* **314**, 103-107 (1985).
- Snodgrass, H. R., Dembic, Z., Steinmetz, M. & von Boehmer, H. *Nature* **315**, 232-233 (1985).
- Haars, R. *et al. J. exp. Med.* **164**, 1-24 (1986).
- Chilson, O. P., Boylston, A. W. & Crumpton, M. J. *EMBO J.* **3**, 3239-3245 (1984).
- Fluscher, D. *Eur. J. Immun.* **14**, 748-752 (1984).
- Hubbard, S. C., Kranz, D. M., Longmore, G. D., Sitkovsky, M. V. & Eisen, H. N. *Proc. Natn. Acad. Sci. U.S.A.* **83**, 1852-1856 (1986).
- Gaudecker, B. & Schmale, E. M. *Cell Tissue Res.* **151**, 347-363 (1974).
- Sun, T. T., Shih, C. & Green, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2813-2817 (1979).
- Rubenfeld, M. R. *et al. J. invest. Derm.* **77**, 221-224 (1981).
- Chu, A. C. *et al. J. invest. Derm.* **81**, 194-197 (1983).
- Hayday, A. C. *et al. Cell* **40**, 259-269 (1985).
- Tutt, M. M. *et al. J. Immun.* **137**, 2998-3001 (1986).
- van den Elsen, P., Shepley, B.-A., Cho, M. & Terhorst, C. *Nature* **314**, 542-544 (1985).
- Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303-2308 (1976).
- White, B. A. & Bancroft, F. C. *J. biol. Chem.* **257**, 8569-8572 (1982).
- Yaketa, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 68-72 (1985).
- Miller, J. *et al. J. Immun.* **134**, 4212-4217 (1985).
- König, F. *et al. Science* **236**, 834-836 (1987).

A new member of the immunoglobulin superfamily—CTLA-4

Jean-François Brunet, François Denizot,
Marie-Françoise Luciani, Magali Roux-Dosseto*‡,
Marie Suzan, Marie-Geneviève Mattei†
& Pierre Golstein

Centre d'Immunologie INSERM-CNRS de Marseille-Luminy,
Case 906, 13288 Marseille Cedex 9, France

* INSERM U.119, 27, Boulevard Leï Roure, 13009 Marseille, France

‡ INSERM U.242, CHU Timone, 13385 Marseille Cedex 5, France

† Present address: UA CNRS 1175, Faculté de Médecine Secteur Nord,
13326 Marseille Cedex 15, France

The immunoglobulin superfamily is a group of proteins, each made of one or several domains sharing key structural features with either the variable (V) or the constant (C) immunoglobulin domains^{1,2}. It includes such functionally important members as the immunoglobulins themselves, major histocompatibility complex (MHC) class I and class II and T-cell receptor (TCR) molecules. Several members of this superfamily are expressed on lymphocytes where they are membrane-bound and capable of interactions with other members of the family, thus taking part in cell-cell recognition. In screening mouse cytolytic-T-cell-derived cDNA libraries, we came across cDNA clones defining a sequence, CTLA-4, which could encode a 223-amino-acid protein clearly belonging to the immunoglobulin superfamily. It consists of one V-like domain flanked by two hydrophobic regions, one of which has a structure suggestive of membrane anchoring. CTLA-4 is mainly expressed in activated lymphocytes and is coinduced with T-cell-mediated cytotoxicity in inducible models of this process. The mouse *ctla-4* gene maps to band C of chromosome 1.

We previously reported the construction of a subtracted cytolytic-T-cell(Tc)-derived complementary DNA library of about 8,300 clones, from which three distinct cytolytic T-cell-associated sequences, CTLA-1, CTLA-2 and CTLA-3, were iso-

lated by conventional differential screenings³. CTLA-1 and CTLA-3 encode distinct serine-esterases, also detected, using similar approaches, as CCP-1 (ref. 4) and H factor⁵, respectively; CTLA-2 encodes a previously undescribed protein product (F.D. *et al.*, in preparation). Further screening of the same subtracted library with a subtracted (Tc-B) cDNA probe allowed the detection of 113 additional clones which had not detectably hybridized with a total Tc-derived cDNA probe. These clones were subjected to successive rounds of screening with several other Tc cDNA probes subtracted against messenger RNA derived from either the B lymphoma M12.4.1, or the EL4 thymoma, or thymocytes. The positive, thus still apparently Tc-specific, clones were then depleted of copies of CTLA-1, -2 and -3 by probing with a pool of inserts representative of these families. The inserts of the 17 cDNA clones left at that stage were used as probes on Northern blots of a panel of RNAs. Only one clone, M17G7, survived this ultimate screening for Tc specificity in that it detected a 2-kilobase (kb) transcript in RNA from the original KB5C20 Tc clone and not from other, non-Tc cells. The M17G7 cDNA insert defines CTLA-4.

A study of the tissue distribution of the CTLA-4 2-kb transcript (Fig. 1) revealed its presence in each of the Tc or Tc-containing populations tested (two clones: KB5C20 and A15.1.17, albeit at a low level for the latter; and three *in vitro* or *in vivo* primary populations: mixed leucocyte culture cells, concanavalin-A-activated blasts and allo-sensitized peritoneal exudate lymphocytes). In addition, like the CTLA-1, -2 and -3 transcripts³, the CTLA-4 transcript was coinduced with cytotoxicity in two systems: on incubation of the constitutively growing cytolytic hybridoma PC60 (ref. 6) with interleukin-containing supernatants and on incubation of concanavalin-A-pulsed thymocytes with an interleukin-2-containing supernatant. A weaker signal was detected in lipopolysaccharide(LPS)-activated blasts (the significance of which is unclear due to the likely heterogeneity of this population) and non-induced PC60 cells, and a very weak one in non-activated thymocytes (Fig. 1) and in non-experimentally-activated peripheral lymphocytes, that is column-purified or unpurified normal spleen cells (data not shown). EL4, the T helper hybridoma T14-117, NK-cell-

Fig. 1 Specificity pattern of CTLA-4 RNA transcripts. Northern blots were prepared with RNA from the following cells: KB5C20, a B10.BR anti-H-2K^b Tc clone¹⁶; A15.1.17, an A.Th anti-A.TL Tc clone¹⁷; MLC cells, A/J spleen cells stimulated for five days *in vitro* with C57BL/6 irradiated spleen cells; Con A blasts, a mixture of C57BL/6 spleen cells stimulated for two or three days *in vitro* with concanavalin A (1.5 $\mu\text{g ml}^{-1}$); PEL, nylon-wool purified BALB/c anti-EL4 peritoneal exudate cells¹⁸; NK, nylon-wool purified Swiss *nu/nu* spleen cells; thymocytes, from C57BL/6 mice; thymocytes plus Con A, the latter incubated for 24 h with Con A (5 $\mu\text{g ml}^{-1}$); thymocytes plus Con A and IL-2, the latter further incubated for five days in the presence of IL-2-containing supernatant from PMA-stimulated EL4 C116 cells; PC60, a mouse $\times$ rat hybridoma⁶; PC60ind, the latter induced with interleukin-containing supernatants¹⁹; EL4, a C57BL/6 thymoma; T14.117, a hybridoma between a T-helper cell clone and the BW5147 thymoma²⁰; IC-21, PU5 and P388D1, macrophage cell lines; PU-5ind, PU-5 activated by a 24-h incubation with supernatant with supernatant from restimulated KB5C20 cells; AtT20, a neurosecretory cell line; M12.4.1, a B lymphoma²¹; LPS blasts, C57BL/6 spleen cells grown for three days with lipopolysaccharide (5 $\mu\text{g ml}^{-1}$); L cells, a C3H fibroblast line; brain and liver cells from C57BL/6 mice. The Northern blots (with about the same amount of RNA per lane according to ethidium bromide staining of gels, not shown) were probed with the radiolabelled 300-bp insert from the CTLA-4 cDNA clone M17G7, isolated from a first, subtracted, KB5C20 cDNA library³. Positive lanes show a main band corresponding to a transcript of about 2 kb, but often also higher bands corresponding to longer transcripts. In parallel, the same cells had been tested for cytotoxicity on EL4, RDM4 and YAC target cells in the absence or in the presence of Con A (10 $\mu\text{g ml}^{-1}$); those found positive in a 4-h or in a 20-h ⁵¹Cr release test are labelled ● or ○, respectively; the others were negative. All methods were exactly as described before³.

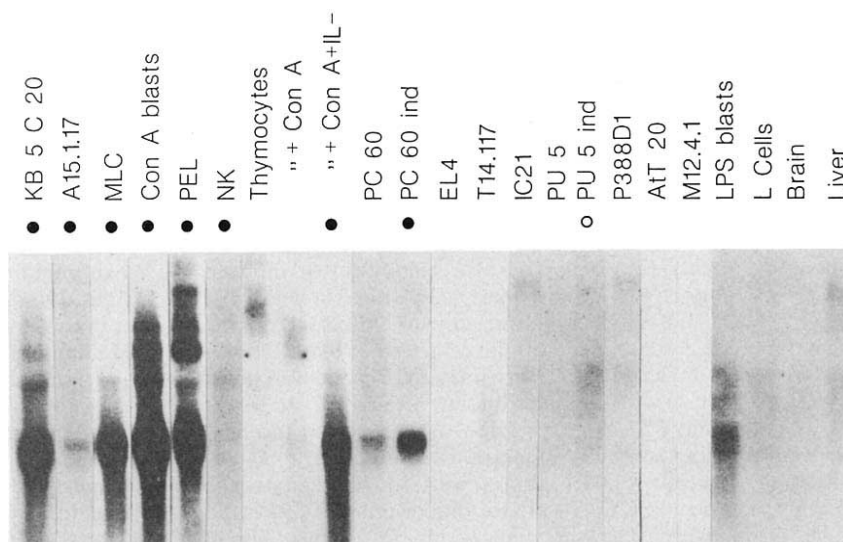
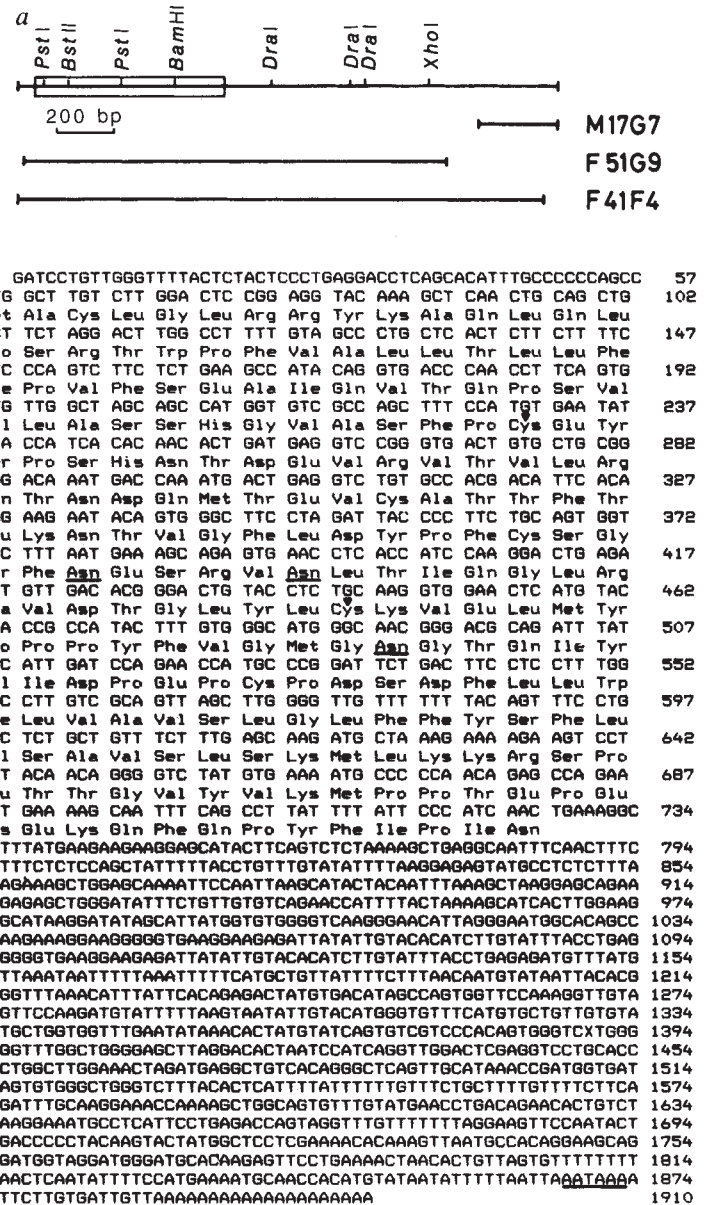


Fig. 2 a, Restriction map of CTLA-4 clones F41F4, F51G9 and M17G7 with restriction enzyme sites used for sequencing. Box: open reading frame. **b**, Nucleotide sequence of CTLA-4 with its predicted amino-acid translation. The nucleotide sequence includes, in its 5' region and for both F41F4 and F51G9 inserts, an in-frame TGA stop codon 30 residues upstream of the first ATG; the latter is part of a GCCATGG consensus sequence²² for initiation of translation; the open reading frame, from nucleotide residue 58 to 726, is followed with a 1.16-kb long untranslated region; at its 3' end, a polyadenylation signal (underlined) precedes a poly(A) stretch. Because of the uncertainties as to cleavage of the leader sequence (see text), the amino-acid sequence is numbered provisionally starting from the initial methionine residue of the conceptual sequence. The latter includes, from 5' to 3': a complex leader sequence (with, first, a few hydrophobic residues, the basic hydrophilic Arg-Arg-Tyr-Lys stretch at positions 7–10, and a long hydrophobic stretch); a possible first residue for the putative mature protein (the glutamine at position 36); the V-like domain (see Fig. 3b and c) with its two key cysteines (at positions 58 and 129) and three Asn-linked (Asn-X-Ser/Thr, where X is any amino acid) potential N-glycosylation sites (underlined); a long hydrophobic, probably transmembrane stretch (residues 162–187); the basic stretch Lys-Met-Leu-Lys-Lys-Arg (residues 188–193); and the essentially hydrophilic 3' end of the protein sequence. Curiously, a seven-residue stretch (residues 213–219) at the end of this cytoplasmic hydrophilic tail of mouse CTLA-4 is also found at residues 126–132 of the polypeptide of human adult T-cell leukaemia virus²³; it will be interesting to find out whether this homology between a cellular gene transcribed in activated T cells and a virus causing T-cell leukaemia is more marked in the human equivalent of CTLA-4. Note that five cysteines are present in the central segment of CTLA-4; of the four belonging to the V-like domain, two can form the disulphide bond critical to the immunoglobulin fold and the other two can form another disulphide bond within this fold. The fifth cysteine, nearest to the putative transmembrane region, might be involved in forming dimeric structures, not unlike the situation in CD8 (ref. 24).

Methods. A second KB5C20-derived cDNA library was constructed according to Gubler and Hoffman²⁵ modified as described³ except that it was size selected in the 1.6–2.2 kb range before it annealed to the pUC9 vector. The library was screened with the 300-bp insert of M17G7 which detected F41F4. The *Pst*I fragment of F41F4 was in turn used as a probe and detected clone F51G9. Sequencing was performed using a chemical cleavage method^{26,27} on both strands of each of two clones except for stretches of the 3' untranslated region.

containing nude spleen cells and also uninduced or induced macrophage lines (Fig. 1) and mast cells (not shown) were negative for the CTLA-4 transcript. Outside of the haematopoietic compartment, all the cells tested (liver, L fibroblasts, brain and the neurosecretory line AtT20) were also negative (Fig. 1). In view of this distribution a tentative statement can be made that CTLA-4 expression is restricted to the lymphoid lineage and is mainly found in activated cells.

Probing a second cDNA library with M17G7 led to the isolation of two more CTLA-4 clones with inserts of 1.85 kb and 1.5 kb respectively. The sequence of these clones (Fig. 2) revealed an open reading frame of 669 nucleotides encoding a putative protein of 223 amino acids. A search for similar sequences at the protein level in the NBRF sequence data bank showed a 22% overall similarity to the human V α ^{VII} domain. The similarity was mainly around the two cysteine residues engaged in the disulphide bond crucial for the immunoglobulin fold. Sequence comparison was therefore extended to other members of the immunoglobulin superfamily. Most of the canonical features of



V-like domains were indeed present (Fig. 3b and c) between residue 39 and residue 140 of CTLA-4. The main discrepancy was the absence of the tryptophan residue at position 52 (in the IgH numbering), which is also the case for instance for Thy-1. A local similarity of up to 70% was found between residue 116 and residue 131 with the third domain of the rabbit polyimmunoglobulin receptor around the corresponding cysteine, where several non-consensus residues are shared between both molecules.

The predicted protein sequence (Fig. 2) and its hydrophobicity pattern (Fig. 3a) reveal that CTLA-4 comprises the V-like domain described above and two flanking regions (Fig. 3b). The 3' flanking region includes, in succession, a 26-residue-long hydrophobic stretch, very probably transmembrane (such hydrophobic regions are, however, replaced with another system of membrane anchoring in at least two members of the immunoglobulin superfamily, Thy-1 ref. 7 and N-CAM 120, ref. 8); the basic stretch Lys-Met-Leu-Lys-Lys-Arg (positions 188–193) resembling that found at the transmembrane-cyto-

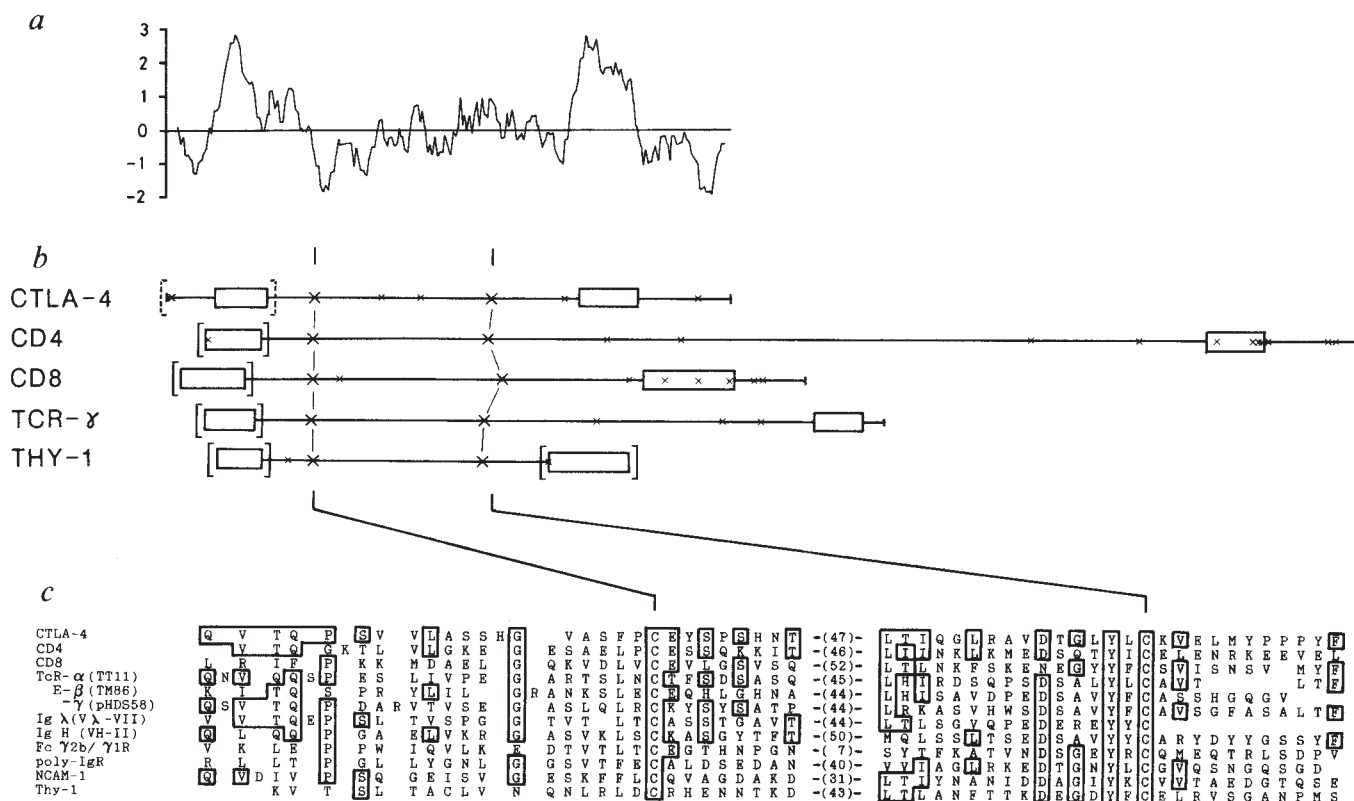


Fig. 3. *a*, Hydropathicity plot²⁸ of CTLA-4, with a window of 11 residues. *b*, Aligned overall primary protein structures of CTLA-4, CD4 (L3T4), CD8 (Lyt-2), γ -chain of the T-cell receptor (TcR) and Thy-1. Boxes, long hydrophobic residues; brackets, peptides known to be cleaved off the mature proteins; stippled brackets, putative CTLA-4 leader sequence; crosses, cysteine residues (note the two unusual extra cysteines in the V-like domain of CTLA-4). *c*, Alignment of the sequences, around the two cysteines involved in the immunoglobulin fold, of CTLA-4, CD4 (ref. 29), CD8 (ref. 30), T-cell receptor α -chain³¹, β -chain³² and γ -chain³³, human V λ chain³⁴ (Ig λ), mouse V μ chain³⁵ (Ig H), Fc γ receptor^{14,36} (Fc γ 2b/ γ 1R), polyimmunoglobulin (Ig) receptor³⁷, the first domain of murine N-CAM³⁸ and Thy-1 (ref. 1). Boxed residues are identified in CTLA-4 and three or more members of the immunoglobulin superfamily (conservative changes like Phe 56 are not boxed).

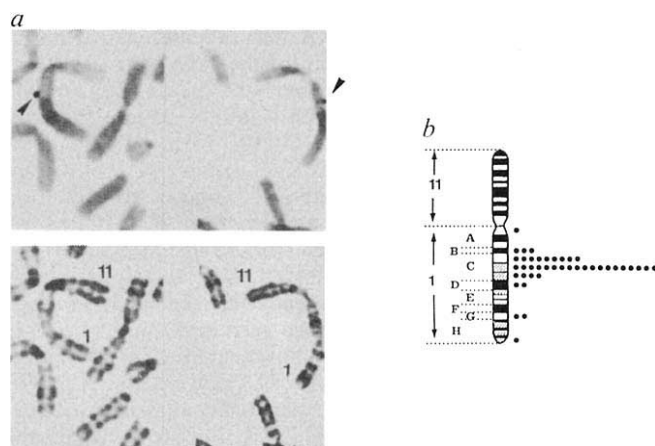


Fig. 4 Localization of the *ctla-4* gene to mouse chromosome 1 by *in situ* hybridization with the CTLA-4 probe F41F4. *a*, Two partial WPM mouse metaphases, showing the specific site of hybridization to chromosome 1. Top, arrowheads indicate silver grains on Giemsa-stained chromosomes, after autoradiography. Bottom, chromosomes with silver grains were subsequently identified by R-banding. Note the homogeneously staining region³⁹ in the C-D band interface of chromosome 1. *b*, Diagram of WPM mouse Rb(1; 11) chromosome, indicating the distribution of labelled sites. In the 123 metaphases examined, there were 250 silver grains associated with chromosomes and 41 of these (16.4%) were localized on chromosome 1. On this chromosome, 78% of the grains mapped to the C band (for banding nomenclature, see ref. 40), with a maximum in the middle part of this band.

Methods. *In situ* hybridizations with probe F41F4 were done using metaphase spreads from a male WPM mouse in which all the autosomes (except 19) are in the form of robertsonian translocations (given by J. L. Guénet, Pasteur Institute). All methods were as indicated before^{3,41,42}. The specific activity of the probe was 6×10^7 d.p.m. μg^{-1} , its final concentration was 25 ng ml^{-1} , and the slides were exposed for 18 days at 4°C .

plasmic hinge of many membrane receptors⁹; and the rest of a hydrophilic putative cytoplasmic tail. The 5' flanking region has the location and hydrophobicity of a leader sequence, but is unusual in its overall length and also perhaps in the basic Arg-Arg-Tyr-Lys stretch at positions 7-10. This leader sequence could be cleaved off between the Ser at position 35 and the Glu at position 36 (taking into account the CTLA-4 hydrophobicity pattern and the site of cleavage for other members of the

immunoglobulin superfamily); alternatively, it may stay uncleaved¹⁰. Antibodies should be of help in determining the location and mode of insertion into the membrane, the precise tissue distribution and possibly the functional (maybe interactive) role of the CTLA-4 protein product.

In a preliminary study of the corresponding gene(s), the Southern blot patterns (not shown) were consistent with there being only one copy of the *ctla-4* gene per mouse haploid

genome. Multiplicity of gene copies cannot therefore account for the several CTLA-4 transcripts longer than the 2-kb one, detected in Northern blots of, in particular, concanavalin A blasts and peritoneal exudate lymphocytes (Fig. 1). In agreement with the Southern blot hybridization patterns, by *in situ* hybridization only one main location was found for the *ctla-4* gene, on band C of mouse chromosome 1 (Fig. 4). The genes for the IgG1/IgG2b FcR-Ly-17 alloantigen and for M1s map apparently more distally on the same chromosome¹¹⁻¹⁴.

Altogether, compared to other currently known members¹⁵ of the immunoglobulin superfamily, the CTLA-4 protein product has in common with Thy-1, CD4 (L3T4), CD8 (Lyt 2) and CD3 (T3) ϵ -chain a single V-like domain most probably expressed at the cell membrane in certain lymphocyte subpopulations. CTLA-4, however, is peculiar in that it is mainly found in activated lymphocytes. The regulation of transcription of the *ctla-4* gene, and also the function of the CTLA-4 protein, seem particularly interesting areas for future studies.

We thank colleagues, especially C. Goridis, B. Malissen, M. Pierres and A.-M. Schmitt-Verhulst at the Centre d'Immunologie, for cells, for advice and for reading the manuscript; B. Jacq (LGBC, Marseille) for considerable help with computer analyses and C. Benoist (Strasbourg) and F. Rougeon (Paris) for advice on cDNA subtraction. Support was through institutional grants from INSERM and CNRS and through grants from the Association pour la Recherche contre le Cancer, the Fondation pour la Recherche Médicale and the Ligue Française contre le Cancer.

Received 1 April; accepted 22 May 1987.

1. Williams, A. F. & Gagnon, J. *Science* **216**, 696-703 (1982).
2. Hood, L., Kronenberg, M. & Hunkapiller, T. *Cell* **40**, 225-229 (1985).
3. Brunet, J.-F. *et al.* *Nature* **322**, 268-271 (1986).

4. Lobe, C. G., Finlay, B. B., Paranchych, W., Paetkau, V. H. & Bleackley, R. C. *Science* **232**, 858-861 (1986).
5. Gershenfeld, H. K. & Weissman, I. L. *Science* **232**, 854-858 (1986).
6. Conzelmann, A., Cortésy, P., Gianfriglia, M., Silva, A. & Nabholz, M. *Nature* **298**, 170-172 (1982).
7. Tse, A. G. D., Barclay, N. A., Watts, A. & Williams, A. F. *Science* **230**, 1003-1008 (1985).
8. He, H.-T., Barbet, J., Chaix, J.-C. & Goridis, C. *EMBO J.* **5**, 2489-2494 (1986).
9. Yarden, Y. *et al.* *Nature* **323**, 226-232 (1986).
10. Braell, W. A. & Lodish, H. F. *Cell* **28**, 23-31 (1982).
11. Kozak, C. A., Davidson, W. F. & Morse, H. C. III *Immunogenetics* **19**, 163-168 (1984).
12. Hibbs, M. L., Hogarth, P. M. & McKenzie, I. F. C. *Immunogenetics* **22**, 335-348 (1985).
13. Holmes, K. L., Palfree, R. G. E., Hämmerling, U. & Morse, H. C. III *Proc. natn. Acad. Sci. U.S.A.* **82**, 7706-7710 (1985).
14. Lewis, V. A., Koch, T., Plutner, H. & Mellman, I. *Nature* **324**, 372-375 (1986).
15. Hunkapiller, T. & Hood, L. *Nature* **323**, 15-16 (1986).
16. Albert, F., Buferne, M., Boyer, C. & Schmitt-Verhulst, A.-M. *Immunogenetics* **16**, 533-549 (1982).
17. Pierres, A., Schmitt-Verhulst, A.-M., Buferne, M., Golstein, P. & Pierres, M. *Scand. J. Immun.* **15**, 619-625 (1982).
18. Berke, G., Sullivan, K. A. & Amos, B. J. *exp. Med.* **135**, 1334-1350 (1972).
19. Erard, F. *et al.* *J. exp. Med.* **160**, 584-599 (1984).
20. Pont, S. *et al.* *Eur. J. Immun.* **15**, 1222-1228 (1985).
21. Hamano, T., Kim, K. J., Leiserson, W. M. & Asofsky, R. J. *Immun.* **129**, 1403-1406 (1982).
22. Kozak, M. *Nucleic Acids Res.* **12**, 857-871 (1984).
23. Seiki, M., Hattori, S., Harayama, Y. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3618-3622 (1983).
24. Johnson, P. & Williams, A. F. *Nature* **323**, 74-76 (1986).
25. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
26. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
27. Chuvpilo, S. A. & Kravchenko, V. V. *FEBS Lett.* **179**, 34-36 (1984).
28. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
29. Tourville, B., Gorman, S. D., Field, E. H., Hunkapiller, T. & Parnes, J. R. *Science* **234**, 610-614 (1986).
30. Nakauchi, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 5126-5130 (1985).
31. Arden, B., Klotz, J. L., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).
32. Hedrick, S. M., Nielsen, E. A., Kovaler, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
33. Saito, H. *et al.* *Nature* **312**, 36-40 (1984).
34. Anderson, M. L. M., Szajnert, M. F., Kaplan, J. C., McColl, L. & Young, B. D. *Nucleic Acids Res.* **12**, 6647-6661 (1984).
35. Bothwell, A. L. M. *et al.* *Cell* **24**, 625-637 (1981).
36. Ravetch, J. V. *et al.* *Science* **234**, 718-724 (1986).
37. Mostov, K. E., Friedlander, M. & Blobel, G. *Nature* **308**, 37-43 (1984).
38. Barthels, D. *et al.* *EMBO J.* (in the press).
39. Traut, W., Winking, H. & Adolph, S. *Cytogenet. Cell Genet.* **38**, 290-297 (1984).
40. Nesbitt, M. N. & Francke, U. *Chromosoma* **41**, 145-158 (1973).
41. Mattéi, M. G. *et al.* *Hum. Genet.* **69**, 268-271 (1985).
42. Camargo, M. & Cervinka, J. *Am. J. hum. Genet.* **34**, 757-780 (1982).

Recombinant human lipocortin 1 inhibits thromboxane release from guinea-pig isolated perfused lung

G. Cirino[†], R. J. Flower[‡], J. L. Browning*, L. K. Sinclair* & R. B. Pepinsky*

School of Pharmacy and Pharmacology, University of Bath, Claverton Down, Bath BA2 7AY, UK

* Biogen Research Corporation, 14 Cambridge Center, Cambridge, Massachusetts 02142, USA

The guinea-pig perfused isolated lung, used in conjunction with the cascade superfusion system to measure the release of thromboxane A₂ (TXA₂), is a simple and convenient model for assessing the inhibition by glucocorticoids of eicosanoid formation¹⁻⁸. Dexamethasone inhibits the release of TXA₂ from the lung when it is stimulated by agents such as RCS-RF² of leukotrienes,^{4,8} but not when bradykinin or arachidonic acid are used¹⁻⁷. Using this model we have shown that the glucocorticoids suppress eicosanoid generation by cells through the induction of a family of phospholipase A₂-inhibitory proteins^{5,6} now termed the 'lipocortins'⁹. Recently the primary structure of one form of lipocortin has been elucidated and the human gene cloned¹⁰. Lipocortin 1 is a polar monomeric protein with anti-phospholipase properties *in vitro*^{10,11} and we now report that when infused into guinea-pig lung preparations this protein has the same inhibitory profile as the

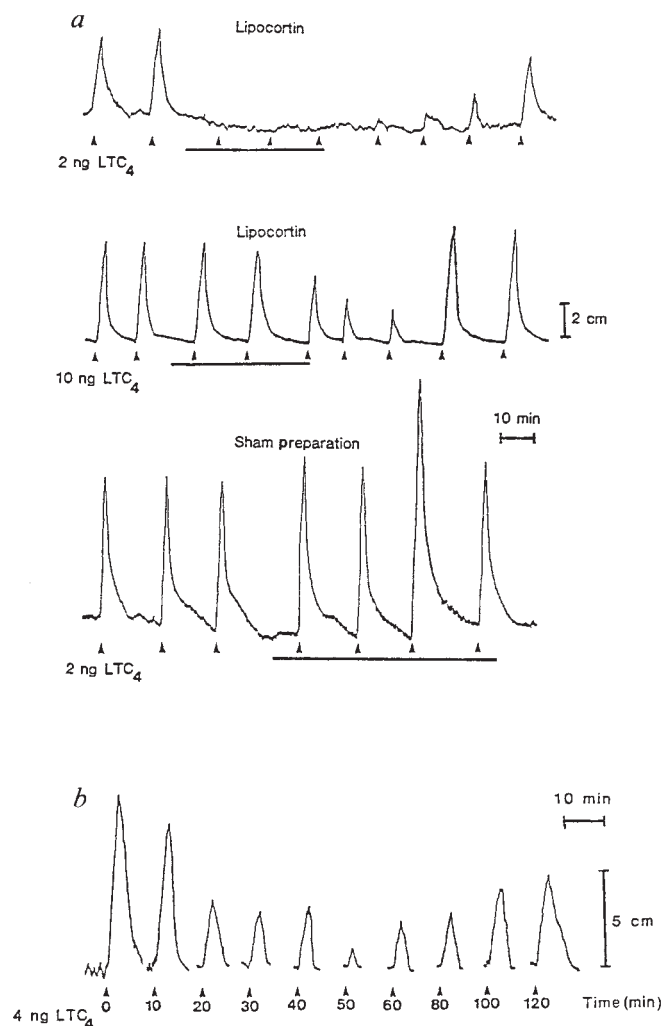
glucocorticoids but with a more rapid onset of action. This is the first demonstration that eicosanoid formation can be inhibited by a recombinant phospholipase inhibitory protein applied extracellularly

Bolus injections of leukotriene C₄ (LTC₄), 1-10 ng, elicited a dose-dependent release of TXA₂ from the guinea-pig perfused lungs (equivalent to ~10-50 ng of the TXA₂ agonist U44619) but had no direct action on the isolated tissues. In 32 out of 32 experiments, in which lipocortin was infused into the lung at concentrations of 0.4-4.0 µg ml⁻¹ (10⁻⁸ M to 10⁻⁷), inhibition of TXA₂ release was detected. The latency and final degree of inhibition achieved depended upon the concentration of lipocortin, the duration of infusion and the dose of LTC₄. Table 1 summarizes mean data obtained from 18 matched experiments, in which different combinations of LTC₄ and lipocortin were tested. Despite variations in potency between batches of recombinant lipocortin, a general dose-response trend is apparent. The top two tracings in Fig. 1a show examples from two experiments where low and high doses of LTC₄ were used. Lipocortin at 0.4 µg ml⁻¹ immediately and completely abolished the release of TXA₂ triggered with 2 ng of LTC₄. Recovery was complete ~70 min after infusion was terminated. With 10 ng of LTC₄, a more graded type of inhibition was observed. In the latter case infusions of 0.4 µg ml⁻¹ lipocortin produced inhibition, but this was not as extensive in magnitude and was slower in onset.

In 4 experiments, a sham preparation of lipocortin was tested against 2, 3 and 10 ng of LTC₄ (see Table 1 with Fig. 1a). In no case was inhibition of more than 10% observed with the sham preparation (this is within the normal variation of the tissue response) and where small effects were seen, they were not sustained (Fig. 1a). A number of other controls were examined. Boiled lipocortin was without effect. Similarly, 40 min infusions of 10 µg ml⁻¹ BSA (Sigma) or 1 µg ml⁻¹ superoxide dismutase (Sigma, bovine kidney) were negative. *Escherichia*

[†] Permanent address: Dipartimento di Farmacologia Sperimentale, Via Leopoldo Rodino 22, 80138, Napoli, Italy.

[‡] To whom correspondence should be addressed.



coli endotoxin (Sigma, serotype 026:B6) at concentrations similar to those contained in the recombinant lipocortin preparations were without effect, but at 1 $\mu\text{g ml}^{-1}$ a transient inhibition was observed. A bolus injection of 10 ng of LTC₄ that had been pretreated with lipocortin (1.0 μg of lipocortin in 1 ml at 20 °C for 10 min) was unimpaired in its ability to induce TXA₂ release indicating that lipocortin itself does not degrade LTC₄. Similarly, no effect (<10% inhibition in six experiments) of lipocortin proteolysed with insoluble trypsin or pronase was observed during a 40-min infusion. When lipocortin (0.4 $\mu\text{g ml}^{-1}$) was infused directly over the assay tissues, it did not alter the sensitivity of the isolated organs to TXA₂ released from the lung by LTC₄; thus lipocortin is not acting as a TXA₂ receptor blocking agent.

Despite the fact that lipocortin was >99% pure, preparations still contained stimulatory substances and the lipocortin itself was present as a mixture of several disulphide-bonded forms. To resolve these further and to enrich the active species, recombinant lipocortin was carried through either of two additional steps, high resolution cation exchange (Mono S Fast Protein Liquid Chromatography, Pharmacia FPLC) and hydrophobic interaction (phenyl-superose FPLC) chromatography. Biological activity was found only in the lipocortin containing fractions.

Fig. 1 a, Inhibition by recombinant human lipocortin 1 of LTC₄-induced TXA₂ generation by the guinea-pig perfused lung. The superfused rabbit aortic strip was used as the assay organ and repetitive bolus injections of LTC₄ (at the points indicated by the arrows) were the stimuli for TXA₂ release. Lipocortin was infused for 40 min into the lung during the period indicated by the bar line. Top, stimulus, 2 ng of LTC₄; lipocortin 400 ng ml⁻¹ (1×10^{-8} M), middle, stimulus, 10 ng of LTC₄; lipocortin 400 ng ml⁻¹ and bottom, stimulus, 2 ng LTC₄; sham lipocortin preparation was infused at a concentration equivalent to 4 $\mu\text{g ml}^{-1}$ of lipocortin. b, Inhibition by lipocortin 1 of LTC₄-induced prostanoind release from the guinea-pig perfused lung. The superfused rat stomach strip which is sensitive to a wider range of prostanoinds was the assay organ. The stimulus was LTC₄ (4 ng) and lipocortin was infused at a concentration of 40 ng ml⁻¹.

Methods. Lungs were prepared as described¹⁴. TXA₂ and prostaglandins were quantitated by bioassay, passing the effluent over rabbit aorta and/or rat stomach strips in cascade¹⁴. The drugs used to elicit the response were dissolved in Krebs' solution or normal saline and injected into the pulmonary artery catheter. When a stable response was achieved, lipocortin was infused. The selectivity of the indicator organs was increased by the infusion of a mixture of antagonists to histamine, acetylcholine and 5-hydroxytryptamine together with indomethacin (1 $\mu\text{g ml}^{-1}$) directly over the tissues^{15,16}. Recombinant human lipocortin 1 was produced in *E. coli*¹⁰ and purified (>99%) by procedures analogous to those previously described¹¹. Lipocortin was dissolved in 25 mM Tris-HCl buffer pH 7.7 with 0.1–5 mM EDTA and 0.1 mg ml⁻¹ bovine serum albumin (fatty acid free, Sigma). The highest endotoxin level in these preparations resulted in a final concentration of 30–40 $\mu\text{g ml}^{-1}$ (limulus assay) when lipocortin was added at 400 ng ml⁻¹ in the perfusing solution. Working solutions were made up fresh for each experiment by diluting aliquots of the lipocortin stock solution into an appropriate quantity of Krebs' solution. *E. coli* of the same strain but lacking the lipocortin-containing plasmid was processed in parallel by methods identical to those used to produce the recombinant protein. This solution was used as a control ('sham lipocortin'). The sham preparations contained both qualitatively and quantitatively the same *E. coli* contaminant proteins and the endotoxin levels were similar to the recombinant lipocortin preparation. Other controls included a buffer blank (the buffer used for the final gel exclusion chromatography step) and a boiled lipocortin preparation (100 °C for 10 min). Precipitated protein was removed by centrifugation and the soluble fraction used in the experiments. Lipocortin was inactivated by proteolysis by treatment with either TPCK-treated trypsin-agarose (Sigma) or pronase-CM-cellulose (Sigma) for 12 h at 33 °C. The insoluble enzymes were removed by centrifugation and the activity of the remaining soluble fragments tested. SDS gel analysis showed that >95% of the protein was digested.

Table 1 Inhibition by human recombinant lipocortin of LTC₄-induced TXA₂ release; summary of matched data

n	LTC ₄ (ng)	Lipocortin final concn μg ml ⁻¹	Mean % inhibition at time (min)							
			10	25	40	50	60	80	90	
Recombinant lipocortin-1										
1	2	0.4	100	100	100	90	80	60	20	
2	3	0.4	49	61	69	0	0	0	0	
1	3	2.0	22	50	76	66	57	40	33	
5	10	0.4	10	16	34	29	30	3	0	
2	10	0.8	0	0	62	0	0	0	0	
4	10	1.2	8	15	32	28	5	3	0	
1	10	2.0	28	61	79	71	61	50	0	
2	10	4.0	0	42	68	74	100	58	30	
Sham preparation of lipocortin										
1	2	4.0	0	9	0	0	0	0	0	
1	3	4.0	7	8	0	0	0	0	0	
1	10	4.0	4	0	6	0	0	0	0	
1	10	4.0	0	0	0	—	—	—	—	

Lipocortin was infused in the final concentrations shown for 40 min and the inhibitions calculated from the traces at various time intervals after the start of the infusion.

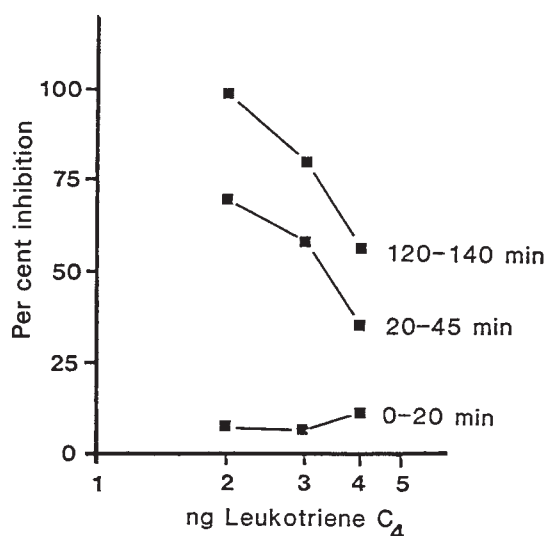


Fig. 2 The cumulative action of recombinant human lipocortin 1 on LTC₄-induced TXA₂ generation by the guinea-pig lung. Three doses of LTC₄ (2, 3 and 4 ng) were given as described in the legend to Fig. 1 during a 2.5-h infusion of 0.4 µg ml⁻¹ lipocortin. The per cent inhibition was calculated with respect to a control dose-response curve constructed from duplicates of each dose given at the beginning of the experiment in the absence of lipocortin.

The Mono S purified lipocortin at a concentration of 40 ng ml⁻¹ in the infusion solution resulted in 80% inhibition of TXA₂ release after 70 min (3 ng LTC₄ stimulation, two experiments), while the phenyl-superoxide purified material inhibited TXA₂ release at levels of 200–400 ng ml⁻¹ (five experiments).

The rat stomach strip can also be used to monitor prostanoid release. Figure 1b shows the release of prostanoids (probably a mixture of TXA₂, PGE₂ and PGF_{2α}) from the guinea-pig lung elicited by repetitive injections of 4 ng of LTC₄. Infusion of 0.4 µg ml⁻¹ lipocortin (40 min) again inhibited prostanoid release and recovery was observed after 80 min. As PGE₂ and PGF_{2α} are detected by this organ, this experiment demonstrates that lipocortin is not working by accelerating the degradation of TXA₂ or by facilitating the conversion of endoperoxides to PGE₂ or PGF_{2α}.

The inhibition of TXA₂ release by lipocortin is a function of time as well as the dose of LTC₄. Figure 2 shows three dose-response curves at 0–20, 25–45 and 50–120 min after commencing infusion of lipocortin at 0.4 µg ml⁻¹. Little change occurred during the first 20 min, but the inhibition increased markedly with continuing exposure to lipocortin.

To ascertain whether the inhibition of TXA₂ release by lipocortin is specific to the agonist LTC₄ we performed experiments using as triggering agents the chemotactic tripeptide f-Met-Leu-Phe, arachidonic acid and bradykinin. The peptide f-Met-Leu-Phe (5–50 ng) produced a reproducible dose-related release of TXA₂ equivalent to ~10–60 ng of the TXA₂ agonist U44619. We found that an infusion of dexamethasone (1 µM final concentration) produced a characteristic time-dependent inhibition resulting in the complete suppression of the TXA₂ release caused by 40–50 ng f-Met-Leu-Phe after 50–60 min (two experiments). Lipocortin (0.2–0.4 µg ml⁻¹) also inhibited f-Met-Leu-Phe induced TXA₂ release (seven experiments); once again the final inhibition achieved was dependent on the dose of f-Met-Leu-Phe, the concentration of lipocortin and the duration of the infusion. For example, when 10 ng of f-Met-Leu-Phe was used as the stimulus, greater than 90% inhibition was seen after a 40-min infusion (two experiments), but with 20 ng of f-Met-Leu-Phe only transient inhibition occurred at the same time point. When 50 ng of f-Met-Leu-Phe was used, a 70-min infusion of lipocortin was required to produce 80% inhibition (2 experiments). Recovery was evident within 15–50 min after termination of the infusion. In 2 experiments, sham lipocortin failed to inhibit the response and in fact a substantial potentiation was seen (5 ng of f-Met-Leu-Phe).

Injections of arachidonic acid (10–50 µg) or bradykinin (5–20 µg) also induced dose-related release of TXA₂ from the perfused organ. Lipocortin, at the highest dose tested

(1.0 µg ml⁻¹) that produced substantial inhibition of both LTC₄ and f-Met-Leu-Phe-induced release was ineffective in preventing TXA₂ release to these agonists. The inclusion of arachidonic acid in our experiments demonstrates that lipocortin is not interfering with the enzymatic generation, release or metabolism of TXA₂ by this perfused organ system. It is not clear why lipocortin or the steroids are unable to block the release of TXA₂ induced by bradykinin. It is possible that separate pools of phospholipase A₂ exist or that a phospholipase C-linked pathway is involved^{3,7}.

Although the anti-phospholipase activity of recombinant human lipocortin 1 has been verified by direct enzymatic analysis^{10,11}, these experiments provide the first demonstration of an activity of the recombinant protein in a biological system. As some of the actions of the anti-inflammatory steroids can be mimicked by lipocortin, the data suggest that steroids may exert some of their anti-inflammatory activities *in vivo* through the induction of the synthesis and/or release of this protein. Whether lipocortin is directly internalized and manifests its activity inside the cell or acts extracellularly is unknown; however, in the simplest model lipocortin would interact with a released or surface-bound phospholipase A₂.

Recently Glenney and colleagues¹² have demonstrated that calpactin II—a protein with Ca²⁺, phospholipid and actin binding properties—is identical to lipocortin I. They have further proposed¹³ that the inhibition of phospholipase A₂ produced by lipocortin is dependent upon its ability to sequester substrate molecules rather than a direct inhibition of the enzyme itself. The experiments reported in this paper do not allow us to distinguish between these two possibilities.

We thank Dr S Bunting of the Upjohn Company for a generous gift of LTC₄. G.C. was supported by a grant from the Biogen Research Corporation.

Received 23 March; accepted 29 May 1987.

- Gryglewski, R. J. *et al. Prostaglandins* **10**, 343–355 (1975).
- Nijkamp, F. P., Flower, R. J., Moncada, S. & Vane, J. R. *Nature*, **263**, 479–482 (1976).
- Blackwell, G. J., Flower, R. J., Nijkamp, F. P. & Vane, J. R. *Br. J. Pharmacol.* **62**, 79–89 (1978).
- Engineer, D. M., Morris, H. R., Piper, P. J. & Sirois, P. *Br. J. Pharmacol.* **64**, 211–218 (1978).
- Flower, R. & Blackwell, G. J. *Nature* **278**, 456–459 (1979).
- Blackwell, G. J. *et al. Nature* **287**, 147–149 (1980).
- Robinson, C. & Hoult, J. R. S. *Eur. J. Pharmacol.* **64**, 211–218 (1978).
- Piper, P. R. & Samhoun, M. N. *Br. J. Pharmacol.* **77**, 267–275 (1982).
- Di Rosa, M. *et al. Prostaglandins* **28**, 441–442 (1984).
- Wallner, B. P. *et al. Nature* **320**, 77–81 (1986).
- Pepinsky, R. B. *et al. J. Biol. Chem.* **261**, 4239–4246 (1986).
- Glenney, J. R. Jr., Track, B. & Powell, M. A. *J. Cell Biol.* **104**, 503–511 (1987).
- Davidson, F. F., Dennis, E. A., Powell, M. A. & Glenney, J. R. Jr. *J. Biol. Chem.* **262**, 1698–1705 (1987).
- Piper, P. J. & Vane, J. R. *Nature* **233**, 29–35 (1969).
- Palmer, M., Piper, P. J. & Vane, J. R. *Br. J. Pharmacol.* **49**, 226–242 (1973).
- Eckenfels, A. & Vane, J. R. *Br. J. Pharmacol.* **38**, 214–221 (1972).

Mapping the limits of the human pseudoautosomal region and a candidate sequence for the male-determining gene

Catrin A. Pritchard, Paul J. Goodfellow
& Peter N. Goodfellow

Laboratory of Human Molecular Genetics, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

The human Y chromosome is composed of two different parts: a pseudoautosomal region shared with the X chromosome which is responsible for sex chromosome pairing and a Y-specific part that encodes the sex determining gene¹. Previously we have shown that the pseudoautosomal gene *MIC2* only rarely recombines between the sex chromosomes and, based on the elevated recombination rates in the pseudoautosomal region, we predicted that this gene would lie close to the Y-specific region². In this report we describe a test of this prediction using long-range restriction mapping techniques. We conclude that *MIC2* is less than 200 kilobases (kb) away from Y-specific sequences. During these experiments we have identified an HTF island in a position consistent with the proposed location of the human sex determining gene.

A meiotic map of the human pseudoautosomal region has been constructed using DNA probes isolated at random from Y-chromosome genomic libraries and probes corresponding to the gene *MIC2* (refs 2 and 3). Genetic distances in the pseudoautosomal region are strikingly different when measured in male as compared to female meiosis. For example, recombination between the telomeric marker *DXYS14* and *MIC2* is close to 50% in male meiosis, whereas in female meiosis it is one-tenth of this value. Several indirect lines of evidence suggest that the physical size of the pseudoautosomal region is smaller than 5,000 kb (refs 3 and 4), implying that the 2% recombination detected between *MIC2* and Y-specific sequences corresponds to a physical interval of 200 kb.

Most XX males occur because of terminal exchange of sequences between the X and Y chromosomes, resulting in the transfer of the sex determining gene *TDF* to the X chromosome⁵⁻⁷. The differing amounts of Y-derived sequences in the genomes of XX males has enabled the construction of deletion maps of the short arm of the Y chromosome^{8,9}. Such maps place the male sex determining gene(s) in a telomeric position, adjacent to the pseudoautosomal region. We have studied three Y-specific loci, defined by probes 47z (ref. 10), GMGY3 (ref. 11) and 27a (ref. 12), all of which map close to *TDF*. The locus defined by 47z is known to map proximal to *TDF*, and to the GMGY3 and 27a sequences¹². However, the relative order of the latter three markers has not been previously determined.

Pulsed field gel electrophoresis (PFGE), using enzymes that cleave infrequently in the mammalian genome, allows the analysis of DNA fragments up to 2,000 kb (refs 13 and 14). To construct a long-range restriction map across the boundary between the two distinct parts of the Y chromosome, DNA from the human cell line Oxen (49; XYYY) was cleaved with combinations of 'rare-cutter' restriction enzymes in single and double digestions, and analysed with the *MIC2* genomic probe 19B, and the three Y-specific probes (see Fig. 1). The major feature of the results is the detection of shared restriction fragments by the probes 19B, 27a and GMGY3 (Table 1). A map constructed from these data is shown in Fig. 2a. For most enzymes, 47z has a distinct pattern of hybridization, although it appears to recognize the same 250 kb *SalI* band as GMGY3 and 27a. A double digestion with *SalI* and *SfiI* shows that this does not reflect their physical linkage.

Table 1 Restriction fragments that hybridize with Y-chromosome probes

Enzyme	Probe			
	19B	27a	GMGY3	47z
<i>SfiI</i>	20 (140) (280)* (450)*	260 (280)* (450)*	260 (280)* (450)*	390 (610)
<i>SalII</i>	90 (280) (510)	250	250	250
<i>NaeI</i>	50 100 140 (280)* (370)*	230 280* (370)*	230 280* (370)*	(170) (200) 300 360
<i>BssHII</i>	(150) 260*	260*	260*	LM
<i>NarI</i>	(150) 260*	260*	260*	LM
<i>ClaI</i>	(250)* 350*	(250)* 350*	(100) 350* (630)	150 (260) (270)
<i>SacII</i>	(150) 260*	260*	260*	LM (120)
<i>NotI</i>	350*	350*	350*	LM
<i>SfiI/SalII</i>	20 (90) (140)	250	250	10 (190) (250)
<i>SfiI/BssHII</i>	(150) 260*	260*	260*	390 (610)
<i>SfiI/NarI</i>	(150) 260*	260*	260*	390 (610)
<i>SfiI/SacII</i>	(150) 260*	260*	260*	390 (610)
<i>SfiI/ClaI</i>	20 (200)*	180 (200)* (260)	80 100 (260)	130

Restriction fragment sizes are given in kb. LM, hybridization to the region of limiting mobility was observed. To confirm the identity of shared bands (asterisks), the experiments were repeated several times using different DNA preparations and pulse times in each case. Weaker bands are in parentheses. In some cases, these may be due to partial cleavage of restriction sites caused by cytosine methylation at a C and G pair. For 19B, fragments extending into the Y-specific region should be four times more intense in the cell line Oxen than equivalent fragments reaching into the X-specific region. By comparing the 19B hybridization patterns of male and female blood DNA with that of Oxen DNA, the X-chromosome specific *SfiI*, *BssHII*, *SacII* and *NarI* restriction fragments have been identified. These are underlined. The probe 47z detects a homologous sequence on Xq (ref. 10) which may account for the fainter bands. Although GMGY3 does identify homologous sequences on some autosomes and also Yq (ref. 11), the hybridization signals are too faint to be considered here. The probe 27a, by contrast, hybridizes to the Y chromosome alone¹². None of the probes contain cleavage sites for the restriction enzymes used in this analysis.

By showing that *MIC2* and two Y-derived sequences can be found on the same DNA restriction fragments, we have bridged the boundary between the pseudoautosomal and Y-specific regions of the Y chromosome. As a result of this, the X and Y chromosomes should possess distinct patterns of hybridization with the *MIC2* genomic probe, 19B. The PFGE analysis of *BssHII*-digested DNA from male and female blood cells and from the lymphoid cell line Oxen confirms this prediction (Fig. 3a and b). A 260-kb band, which is recognized by 19B, 27a and GMGY3, correlates with the presence of the Y chromosome; it is absent from female blood DNA and has a fourfold increase in intensity in Oxen. The PFGE fragment of 150 kb is present in all cell lines and represents the X-specific band. *NarI* diges-

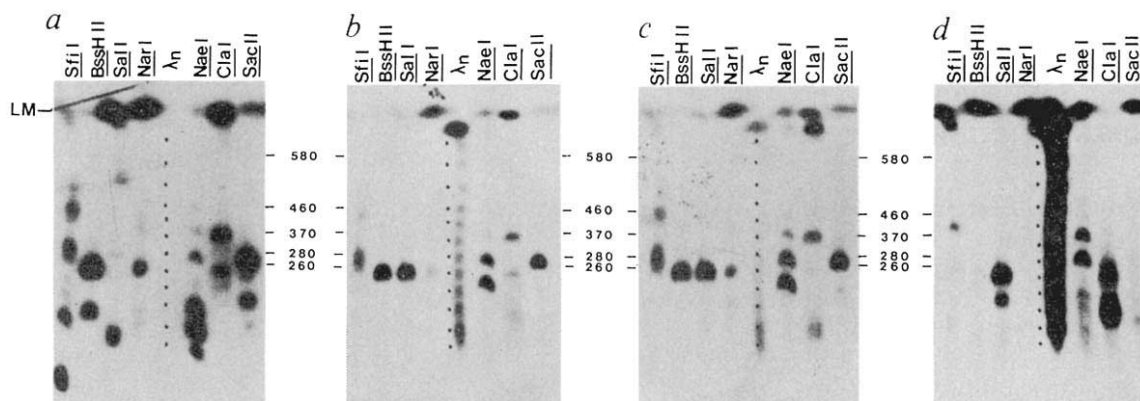
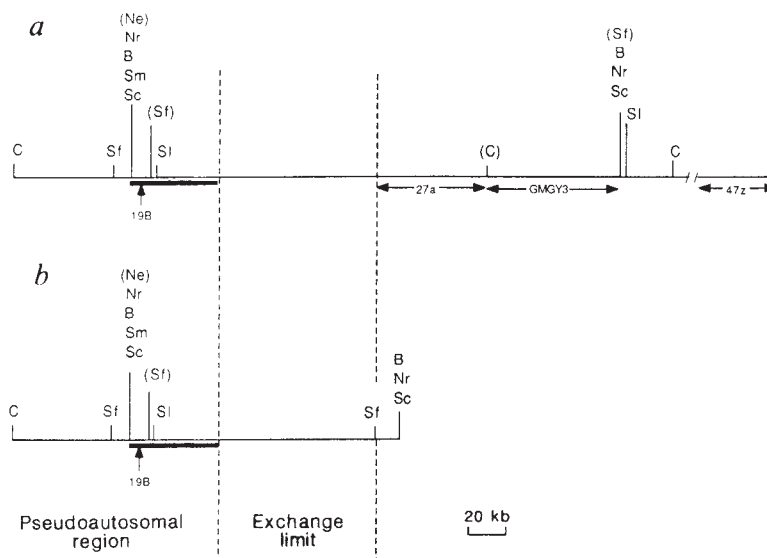


Fig. 1 PFGE analysis of genomic DNA from the human lymphoid cell line Oxe (ref. 20). This cell line contains 49 chromosomes: a complete set of autosomes, one X and four Y chromosomes. The DNA in each track had been cut with the enzymes shown and subjected to PFGE separation for 18 h using a switch time interval of 60 s. After Southern transfer, the DNA was hybridized consecutively with the four probes: a, probe 19B; b, 27a; c, GMGY3; d, 47z. The same filter was used in each case so that comigrating bands could be identified. The black dots in a–d show the positions of multimers of phage λ cI857 which has a unit size of 48.5 kb. In b and d, visualization of the markers was by hybridization to λ DNA, although the autoradiograph of d is considerably over-exposed. Yeast chromosomes provided additional markers and these were separated in the tracks at the ends of the gel (sizes in kb are shown). The rate of migration of the DNA appears to increase towards the outer tracks. Estimates for DNA fragment sizes have therefore been averaged across the gel to the nearest 10 kb by comparison with both yeast and λ markers. Common restriction fragments are observed by the probes 19B, 27a and GMGY3 in *Sfi*I (280 kb and 450 kb), *Bss*HII (260 kb), *Nar*I (260 kb), *Nae*I (280 kb and 370 kb) and *Sac*II (260 kb) digestions. For *Cla*I, a fragment of 350 kb is shared by 19B, 27a and GMGY3. A *Cla*I site that is sensitive to partial cleavage is located within this fragment. When it is cut, it generates two fragments of 250 kb (shared by 19B and 27a) and 100 kb (recognized by GMGY3) respectively. Few bands are observed with the probe 47z. Hybridization to the region of limiting mobility (LM) is seen when the DNA is cut with *Bss*HII, *Nar*I and *Sac*II, representing unresolvable DNA fragments larger than 800 kb.

Methods. High-molecular-weight DNA was prepared in agarose blocks as described in detail elsewhere²¹. Approximately 3×10^5 cells, representing 2 μ g of DNA were encapsulated in each block. The DNA was cut with 20 units of enzyme for 6–8 h as recommended by the manufacturer. After digestion, the blocks were loaded into the slots of a 0.8% agarose gel and electrophoresed at 350 V in a PFGE apparatus (LKB Pulsaphor). The pulse time was as indicated above and the buffer temperature was maintained at 12 °C during the run. The DNA was acid hydrolysed (0.25 M HCl for 15 min) after ethidium bromide staining. Southern transfer to Gene Screen Plus was as recommended by the supplier (NEN). DNA inserts were used as probes and radiolabelled by the random priming method²² to a specific activity of at least 10^8 c.p.m. per μ g DNA. The filters were hybridized in 1 M NaCl, 10% dextran sulphate, 1% SDS at 65 °C and washed the next day in $0.2 \times$ SSC at 65 °C. Autoradiography was carried out at –70 °C for 1–4 days. The filters were stripped of radioactivity by treatment in 0.4 M NaOH for 30 min at 42 °C and then in 0.2 M Tris-base, pH 7.5, $2 \times$ SSC and 0.2% SDS under the same conditions. Care was taken to ensure complete removal of probe before rehybridization.

Fig. 2 Long-range restriction map across the pseudoautosomal region/sex-chromosome-specific region boundary. Restriction enzymes: Sf, *Sfi*I; Sl, *Sal*I; Sc, *Sac*II; C, *Cla*I; B, *Bss*HII; Ne, *Nae*I; Nr, *Nar*I; Sm, *Sma*I. Partially cleaved sites are shown in parentheses. The *MIC2* gene extends over 50 kb and is indicated by the black bar. a, The Y chromosome. The HTF island associated with the *MIC2* gene is known to possess restriction enzyme sites for *Nae*I, *Sma*I, *Sac*II, *Bss*HII and *Nar*I (manuscript in preparation). Double digestions locate this island between the two *Sfi*I sites that are separated by 20 kb. As the 19B probe shares several common PFGE fragments with 27a and GMGY3, it must be located on the proximal side of the island. This view is supported by PFGE studies of *MIC2* genomic sequences known to reside on the opposite side of the island (data not shown). The *Nae*I site in the *MIC2* island is partially cleaved in some cell lines. The relative orientation of the *Nae*I fragments could therefore not be interpreted without ambiguities. Digestions with *Cla*I have established the order *MIC2*–27a–GMGY3 and the GMGY3 locus has been placed within 80 kb of the proximal HTF island. b, The X chromosome. This map is based on the X-chromosome-specific bands observed in Oxe (Table 1) and the PFGE analysis of male and female blood DNA. Comparison of a and b delineates the pseudoautosomal region boundary, indicated by dotted lines.



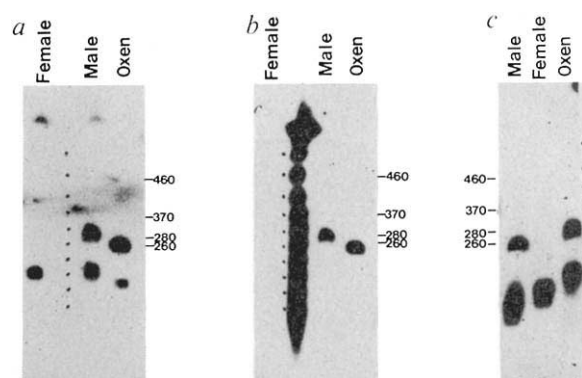


Fig. 3 PFGE analysis of DNA from the Oxen cell line and from female and male blood. The gels were run under the same conditions as in Fig. 1. *a*, Digestions with *Bss*HII. Hybridization with the *MIC2* genomic probe 19B indicates the presence of X- (150 kb) and Y- (260 kb) chromosome-specific bands. The same filter was hybridized in *b*, with the probe 27a which recognizes the Y-derived 260-kb band in the male cell lines but fails to react with female DNA. *c*, Digestions with *Nar*I. Probe 19B detects an X-chromosome-specific band at 150 kb and a Y-specific band at 260 kb. Concatenated λ was used as a marker in the middle lane of the gels: multimers of the 48.5-kb monomer are indicated. Yeast chromosomes were electrophoresed in the outer tracks. Considerable variation in the DNA migration rate is observed across the gel. Fragment sizes have been adjusted to account for this variation, but in some cases the variation may be due to differences in the amount of DNA loaded.

tions also demonstrate the detection of sex-chromosome-specific fragments by the probe 19B (Fig. 3c). Further PFGE analysis has made it possible to define the point at which the X and Y chromosomes are no longer homologous. The exchange limit must lie centromeric to the 3' end of the *MIC2* gene (~50 kb proximal to the 19B sequence; unpublished data) but distal to the X-chromosome-specific *Sfi*I, *Bss*HII and *Nar*I restriction sites that lie within 140–150 kb of the 19B sequence (Fig. 2). This means that in male meiosis in the pseudoautosomal region, 2% recombination occurs in <150 kb of DNA. This rate is 10–20 times higher than the genome average¹⁵.

When Oxen DNA is restricted with *Sfi*I, *Bss*HII, *Sac*II and *Nar*I the PFGE fragments recognized by 27a, 19B and *GMGY3* are all of 260 kb (Fig. 1 and Table 1). Furthermore, double digestions with combinations of these enzymes do not alter the size of the cognate band (Table 1). Based on these results, we suggest that the sites for these enzymes are clustered at both ends of a 260-kb interval. Such a clustering of sites is characteristic of the short unmethylated C- and G-rich regions known as *Hpa*II tiny fragment (HTF) islands. Recent investigations have demonstrated a close association of HTF islands with the 5' ends of expressed genes¹⁶. In support of this, we have found an HTF island at the 5' end of the *MIC2* gene (unpublished observations) and this correlates with the distal cluster of rare restriction sites detected by 19B (Table 1 and Fig. 2). The proximal HTF island is Y-chromosome-specific and close to the locus for *GMGY3*; a *Cla*I-*Sfi*I double digest establishes that the distance between the two is less than 80 kb (Fig. 2a).

We have also identified an X-chromosome-specific HTF island ~150 kb proximal to *MIC2* (Table 1 and Fig. 2b). This HTF island could represent XG, the most distal of the known X-linked genes¹⁷. If this proves to be the case, then the close proximity of *MIC2* and XG supports the *cis*-regulatory model of the relationship between the expression of these two genes¹⁸.

The PFGE analysis presented here establishes the order: telomere-*MIC2*-27a-*GMGY3*-47z-centromere. Although *TDF* cannot be unambiguously placed on the map the most likely position is between the loci defined by *GMGY3* and 47z. In a

series of eighteen XX males, sixteen were found to have inherited both the *GMGY3* and 27a loci; several of the sixteen positive XX males had failed to inherit the 47z locus¹². If it is assumed that the two exceptional XX males are caused by terminal exchange and inheritance of Y-derived sequences then *TDF* would map between *MIC2* and the 27a locus. However, Petit *et al.* have presented convincing evidence to show that these XX males have a different origin not involving X-Y exchanges⁷. This is consistent with the *GMGY3* and 27a sequences mapping distal to *TDF*. That there is an HTF island proximal to the *GMGY3* locus is striking, given the absence of known genes in this region, other than *TDF* (ref. 19).

We thank our colleagues in the Laboratory of Human Molecular Genetics at ICRF for their support and help in setting up the PFGE apparatus, John Bell, Isabelle Henry and Denise Barlow for their advice in the initial stages of this work, Nabeel Affara (Duncan Guthrie Institute, Glasgow) for *GMGY3* and Jean Weissenbach (Institut Pasteur) for 47z.

Received 22 April; accepted 28 May 1987.

- Burgoyne, P. S. *Hum. Genet.* **61**, 85–90 (1982).
- Goodfellow, P. J., Darling, S. M., Thomas, N. S. & Goodfellow, P. N. *Science* **234**, 740–743 (1986).
- Rouyer, F. *et al. Nature* **319**, 291–295 (1986).
- Mondello, C., Ropers, H.-H., Craig, I. W., Tolley, E. & Goodfellow, P. N. *Ann. hum. Genet.* **51**, 137–143 (1987).
- Ferguson-Smith, M. A. *Lancet* **ii**, 475–476 (1966).
- De la Chapelle, A. *Hum. Genet.* **58**, 105–116 (1981).
- Petit, C. *et al. Cell* **49**, 595–602 (1987).
- Vergnaud, G. *et al. Am. J. hum. Genet.* **38**, 109–124 (1986).
- Affara, N. A. *et al. Nucleic Acids Res.* **14**, 5375–5387 (1986).
- Geldwerth, D. *et al. EMBO J.* **4**, 1739–1743 (1985).
- Affara, N. A. *et al. Nucleic Acids Res.* **14**, 5352–5373 (1986).
- Pritchard, C. A., Goodfellow, P. J. & Goodfellow, P. N. *Nucleic Acids Res.* (in the press).
- Schwartz, D. C. & Cantor, C. R. *Cell* **37**, 67–75 (1984).
- Carle, G. F. & Olson, M. V. *Nucleic Acids Res.* **12**, 5647–5664 (1984).
- Renwick, J. H. *Brit. med. Bull.* **25**, 65–73 (1969).
- Brown, W. R. A. & Bird, A. P. *Nature* **322**, 477–481 (1986).
- Geller, R. L., Shapiro, L. J. & Mohandas, T. K. *Am. J. hum. Genet.* **38**, 884–890 (1986).
- Goodfellow, P. J., Pritchard, C., Tippet, P. & Goodfellow, P. N. *Ann. hum. Genet.* **51**, 161–167 (1987).
- Goodfellow, P., Darling, S. & Wolfe, J. J. *med. Genet.* **22**, 329–344 (1985).
- Sirota, L., Zlotogora, Y., Shabtai, F., Halbrecht, I. & Elian, E. *Clin. Genet.* **19**, 87–93 (1981).
- Bernards, A., Koeter, J. M., Michels, P. A. M., Mobergs, R. M. P. & Borst, P. *Gene* **42**, 313–322 (1986).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).

A novel receptor-operated Ca^{2+} -permeable channel activated by ATP in smooth muscle

C. D. Benham* & R. W. Tsien†

Department of Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

Receptor-operated Ca^{2+} entry has been proposed as a signalling mechanism in many cells^{1–8}. Receptor-operated Ca^{2+} channels (ROCs) were first postulated in smooth muscle by Bolton², van Breemen³ and Somlyo and Somlyo⁴, but recordings of directly ligand-gated Ca^{2+} current are lacking. Here we describe receptor-operated Ca^{2+} current evoked in arterial smooth muscle cells by ATP, a sympathetic neurotransmitter⁹. ATP activates channels with approximately 3:1 selectivity for Ca^{2+} over Na^{+} at near-physiological concentrations and with a unitary conductance of ~5 pS in 110 mM Ca^{2+} or Ba^{2+} . The channels can be opened even at very negative potentials and resist inhibition by cadmium or nifedipine, unlike voltage-gated Ca^{2+} channels; they are not blocked by Mg^{2+} , unlike NMDA (*N*-methyl-D-aspartate)-activated

* Present address: Department of Pharmacology, SK&F Research, The Frythe, Welwyn, Hertfordshire, UK.

† To whom correspondence should be addressed.

Fig. 1 Membrane current responses of ear artery smooth muscle cells to extracellular application of ATP. *a*, Membrane current response of a cell under voltage-clamp to bath application of $1 \mu\text{M}$ ATP (solid bar). The holding potential (h.p.) was -50 mV and depolarizing voltage pulses to -30 mV were applied at 0.83 Hz (inset). *b*, Membrane current response of another cell to application of ATP- $\gamma\text{-S}$ ($10 \mu\text{M}$) from a puffer pipette (aperture $5 \mu\text{m}$) placed about $100 \mu\text{m}$ from the cell. Positive pressure of ~ 2 pounds per square inch was applied for the duration of the horizontal bar (1 s). Dashed line, continued presence of the applied solution as judged by control experiments with dye in the puffer pipette. Inset, voltage protocol (h.p. = -60 mV to 0 mV for 190 ms , then to -40 mV for 165 ms), repeated at 0.83 Hz . *c*, Instantaneous current-voltage relationship for ATP-activated inward current. Cells were held at -60 mV and voltage steps were applied before and during ATP application as in *a*. To allow comparison of ATP-induced responses in different cells, currents recorded just after a test step were normalized by the ATP-induced current at -60 mV just before the step. Each point is the average of normalized responses in two to five cells at a given voltage, collected from a total of ten cells.

Methods. Smooth muscle cells were enzymatically dissociated from the central ear arteries of adult NZW rabbits as previously described^{29,20}, plated onto glass coverslips in modified Tyrode's solution containing (mM): NaCl 130, KCl 5, MgCl_2 1.2, CaCl_2 1.5, glucose 10, HEPES 10 (pH adjusted to 7.2 with NaOH). Cells were stored at 4°C , and were used within 24 h of dispersion. Voltage clamp recordings were carried out with the patch-clamp method¹⁵. In the recordings presented in Fig. 1, cells were bathed in modified Tyrode's solution and the pipette (intracellular) solution contained (mM): CsCl 130, MgCl_2 3, EGTA 10, HEPES 5 (pH adjusted to 7.2 with tetraethyl ammonium hydroxide, TEA-OH). All experiments were performed at room temperature ($\sim 22^\circ\text{C}$). Current signals were stored on tape without filtering for later analysis or sampled by digital computer at 5 kHz after filtering with an 8-pole Bessel filter set at 1 kHz (-3 dB).

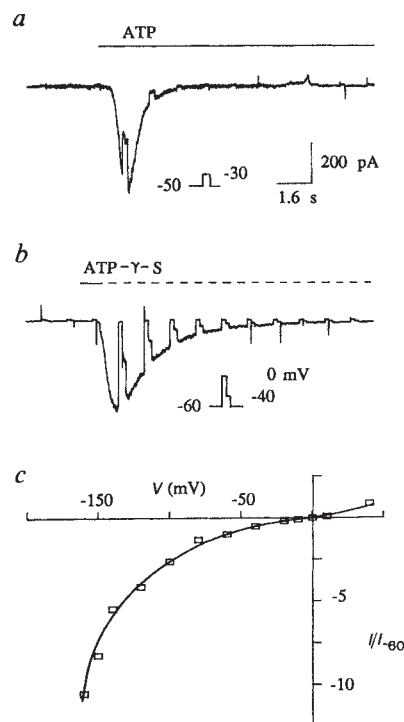
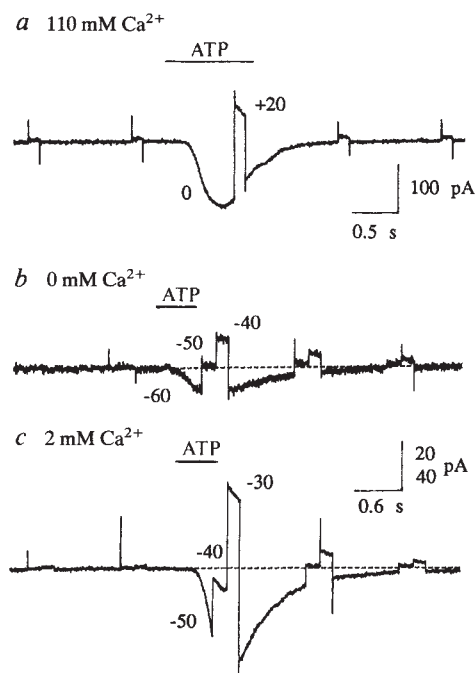


Fig. 2 Responses to pressure application of ATP in different extracellular solutions. *a*, ATP-evoked inward current in a cell bathed in isotonic calcium. Holding potential 0 mV and test potential $+20 \text{ mV}$. Bathing solution contained (in mM): CaCl_2 110, glucose 10, HEPES 10 (pH adjusted to 7.2 with TEA-OH). Pipette solution as in Fig. 1. *b, c*, Effect of 2 mM Ca^{2+} on reversal potential of ATP response in presence of a relatively impermeant extracellular cation and high internal sodium. Internal solution contained (mM): NaCl 130, glucose 10, EGTA 10, histidine 10 (pH adjusted to 7.2 with NaOH). *b*, ATP-activated current in a cell bathed in zero sodium, divalent cation-free solution containing (mM): glucosamine 135, glucose 10, histidine 10 (pH 6.3). At this pH, glucosamine ($pK_a = 8.6$) should be 99% ionized. Responses in Tyrode's solution were unaffected by this pH change. The reversal potential of the ATP-activated current was estimated by interpolation as -50.3 mV . *c*, Response in another cell bathed in sodium-free, glucosamine solution with addition of 2 mM CaCl_2 . The interpolated reversal potential was -38.0 mV .



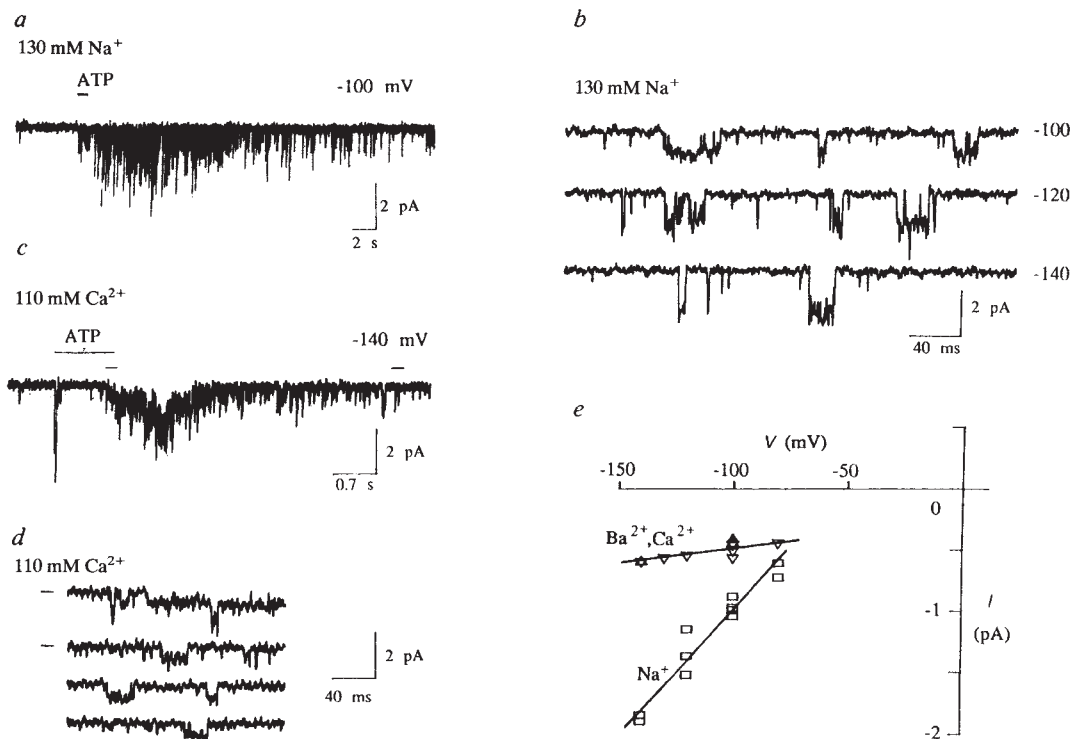
channels^{10,11}; they are directly activated by ligand, without involvement of readily diffusible second messengers, unlike cation channels in neutrophils¹² and T lymphocytes¹³. Thus, the ATP-activated channels provide a distinct mechanism for excitatory synaptic current and Ca^{2+} entry in smooth muscle.

Current recordings were made from freshly dispersed smooth muscle cells from rabbit ear artery¹⁴ with patch-clamp techniques.¹⁵ Figure 1*a* shows a whole-cell recording from a cell held at -50 mV in a standard Tyrode solution. Bath application of $1 \mu\text{M}$ ATP evoked an inward current of several hundred pA,

typical of responses with 130 mM external Na^+ (ref. 14). Desensitization developed in seconds, whereas recovery from desensitization required several minutes following removal of micromolar ATP. Small responses ($10\text{--}20 \text{ pA}$ at -60 mV) to low concentrations of ATP (10^{-8} M) could be obtained without much desensitization with repeated applications every 30 s.

Both free ATP^{4-} and divalent cation-complexed ATP were effective, judging by the potency of the nucleotide in solutions lacking free divalent cations (Fig. 3*a*) or containing isotonic Ca^{2+} (Figs 2*a* and 3*c*) or Mg^{2+} . ATP hydrolysis was not

Fig. 3 Unitary currents activated by ATP in excised outside-out patches. *a*, Response of an outside-out patch bathed in an external solution containing (in mM): 130 Na, 1 EGTA (no added divalent cations), 5 glucose and 10 HEPES (pH 7.2 with NaOH). *b*, Unitary Na^+ currents from the same patch displayed on an expanded time base. Records were collected at three different holding potentials following three consecutive applications of ATP spaced ~ 2 min apart. *c*, *d*, ATP-activated Ca^{2+} current in an outside-out patch, h.p. = -140 mV, external solution contained isotonic Ca^{2+} as in Fig. 2*a*; pipette solution as in Fig. 2*a* but with addition of 1 mM BAPTA. Transient inward current at start of puffer application is a pressure artefact. Latency of ATP-activated inward current largely



reflects delay in the arrival of test solution; the puffer pipette was deliberately pointed away from the patch to avoid rapid desensitization associated with direct application. *d*, Unitary Ca^{2+} currents from the same experiment on an expanded time scale. Mean amplitude of unitary openings is 0.59 pA. The first two traces correspond to the two regions marked by short horizontal bars in *c*, whereas the other two records were taken after the stretch of record displayed in *c*. *e*, Voltage-dependence of unitary current amplitude with bathing solutions containing 110 mM Ca^{2+} (Δ), 110 mM Ba^{2+} (∇) and 130 mM Na^+ ($\square$). Each point is the mean unitary current amplitude for an individual patch. Data from a total of 13 patches. Straight lines are fitted by least-squares criteria.

necessary for the response or its desensitization, because large transient inward currents were also evoked by non-hydrolysable analogues such as ATP- γ -S (Fig. 1*b*) or α, β -methylene ATP (AMPCP-P). ADP gave little or no response at 10 μM and inward currents about fivefold smaller than ATP responses at 0.1 mM (four cells). Adenosine (1 mM) was completely ineffective (three cells). The overall sequence of potency, AMPCP-P = ATP > ADP > adenosine, is characteristic of P_2 purinergic receptors⁹. Previous findings in whole-ear artery¹⁶ show potent contractile effects of AMPCP-P, as in our single cells, but weaker responses to ATP itself, possibly owing to ATP hydrolysis by ectoenzymes.

Current responses to voltage pulses (Fig. 1*a* and *b*) indicate that ATP activates a large increase in conductance with a reversal potential (E_{rev}) close to 0 mV (Fig. 1*c*; ref. 14). The ATP-activated current increased almost linearly between 0 and -60 mV (Fig. 1*c*), and more steeply below -80 mV (see also ref. 17). There was no sign of a decrease in inward current down to -160 mV, despite the presence of 1.2 mM Mg^{2+} in the extracellular solution.

The selectivity of the channels was evaluated with measurements of E_{rev} with various external and internal solutions (Fig. 2, Table 1). Previous results have shown that the ATP-activated conductance rejects Cl^- but shows little selectivity among monovalent cations¹⁴. To determine whether Ca^{2+} is permeant, we carried out whole-cell recordings with Ca^{2+} as the major external cation. With 110 mM CaCl_2 outside, ATP-activated inward currents were observed even with holding potential (h.p.) = 0 mV (Fig. 2*a*) and were as large as 500 pA at h.p. = -60 mV, comparable in size to currents in Tyrode solution with 130 mM Na^+ (Fig. 1*a*). With 110 mM external Ca^{2+} , E_{rev} averaged $+5.7 \pm 0.3$ mV (Table 1), corresponding to a relative selectivity for Ca^{2+} over Cs^+ of 2.7 according to a modified

Goldman-Hodgkin-Katz equation¹⁸. Inward currents were also seen with isotonic (110 mM) Ba^{2+} or Mg^{2+} in the external solution. Reversal potential determinations gave permeability ratios $P_{\text{divalent}}/P_{\text{Cs}}$ in the sequence Ca^{2+} (2.7) > Ba^{2+} (1.9) > Mg^{2+} (0.8).

Additional E_{rev} determinations were made with 130 mM Na^+ in the pipette and various concentrations of Na^+ or Ca^{2+} in the external medium, using glucosamine ion (single positive charge) as a substitute for Na^+ . A significant permeability of glucosamine ($P_{\text{glu}}/P_{\text{Na}} \approx 0.15$) was indicated by reversal potentials obtained with 135 mM glucosamine (-48.7 mV) or 5 mM Na^+ , 135 mM glucosamine (-41.5 mV). With 2 mM Ca^{2+} , 0 mM Na^+ (glucosamine) in the bathing solution, E_{rev} averaged -38.6 mV (Fig. 2*b*), corresponding to a permeability ratio $P_{\text{Ca}}/P_{\text{Na}} = 3.3$. With this relative preference for Ca^{2+} , E_{rev} would be expected to shift by $+1.9$ mV when Ca^{2+} is increased from 0 to 10 mM in the presence of 121 – 135 mM Na^+ . The measured change in reversal potential ($+3$ mV, Table 1) is in reasonable agreement with this prediction. We conclude that ATP-activated channels display moderate selectivity for Ca^{2+} over Na^+ , albeit far less than voltage-gated Ca^{2+} channels¹⁹.

We found additional differences between receptor-operated channels opened by ATP and voltage-gated Ca^{2+} channels with regard to pharmacological responsiveness to Ca^{2+} channel blockers. The ATP-activated current was not reduced by nifedipine (5 μM , two cells), or by cadmium (0.5 mM, three cells), agents known to inhibit voltage-gated Ca^{2+} channels in ear artery cells²⁰. Conversely, we saw no effect of 1 μM ATP on voltage-gated Ca^{2+} channels of either T or L type.

Experiments on membrane patches enabled us to study receptor-channel coupling and single-channel properties. We were particularly interested in determining whether the ATP-activated current might involve a Ca^{2+} -activated, non-selective cation

Table 1 Reversal potentials of ATP-activated current

External solution (mM)	Internal solution (mM)	E_{rev} mean $\pm$ s.e.m.	(n)	$P_{divalent}/P_{monovalent}$
110 Ca^{2+}	130 CS^+	$+5.7 \pm 0.3$	(3)	2.7
110 Ba^{2+}	130 CS^+	$+1.0 \pm 0.9$	(3)	1.9
110 Mg^{2+}	130 CS^+	-13.9 ± 0.5	(5)	0.8
0 Ca^{2+} , 0 Na^+ (135 glu)	130 Na^+	-48.7 ± 1.5	(4)	—
2 Ca^{2+} , (135 glu)	130 Na^+	-37.5 ± 0.6	(4)	3.3
10 Ca^{2+} , (122 glu)	130 Na^+	-27.3 ± 1.6	(5)	2.1
0 Ca^{2+} , 5 Na^+ (135 glu)	130 Na^+	-41.5 ± 1.9	(4)	—
0 Ca^{2+} , 135 Na^+	130 Na^+	$+2.2 \pm 0.3$	(6)	—
10 Ca^{2+} , 121 Na^+	130 Na^+	$+5.3 \pm 0.1$	(5)	5.5

Values of E_{rev} were obtained by interpolation between potentials negative enough or positive enough to give ATP-activated currents that were clearly inward or outward. Junction potentials for solution pairs in rows 1–3 (-6 mV for Ca^{2+} , -7 mV for Mg^{2+} and Ba^{2+} , pipette negative) were corrected for in estimates of E_{rev} . The value of E_{rev} in the eighth row is very close to a theoretically expected value of 1.1 mV, so other voltage offsets were small. External solutions containing glucosamine hydrochloride (glu) were buffered with 10 mM histidine (pH 6.3) so that the glucosamine (pK_a 8.6) was $>99\%$ ionized. In these experiments, pipette (internal) solutions were also buffered with 10 mM histidine (pH 7.2). A value of 0.15 for P_{glu}/P_{Na} was taken into account in calculations of P_{Ca}/P_{Na} when appropriate. Monovalent ion activities were taken as 0.75 and divalent activities as 0.25 (110 mM divalent ion) or 0.3 (2 – 10 mM divalent ion).

conductance^{12,21,22}. Such an explanation seems very unlikely because ATP-activated inward currents were found in over 90% of stable outside-out patches, despite the presence of 10 mM EGTA and 1 mM BAPTA in the pipette (estimated free Ca^{2+} concentration $<10^{-8}$ M). As a test for the involvement of diffusible second messengers, we also recorded from cell-attached patches while applying ATP to the rest of the cell. No inward currents were ever seen in these recordings (five patches from five cells), whereas recordings from outside-out patches from the same batch of cells consistently showed ATP-activated current (three out of three patches).

Unitary currents were resolved in many experiments by adjusting the position of a puffer pipette containing 1 μ M ATP to give a submaximal current response. Figure 3a shows a record from an outside-out patch clamped at h.p. = -100 mV with 130 mM external Na^+ (1 mM EGTA, no added divalent cations). ATP evoked inward currents, including clearly resolved unitary current pulses ranging in amplitude from 1 pA at -100 mV to 2 pA at -140 mV (Fig. 3b and e).

Figure 3c and d shows an outside-out patch recording with Ca^{2+} as the external cation. Here, unitary currents were considerably smaller than with external Na^+ , but were still clearly resolved (Fig. 3c). Collected data from patches with 110 mM Ca^{2+} or 110 mM Ba^{2+} are very similar (Fig. 3e) and yield a unitary conductance of ~ 5 pS, assuming a reversal potential close to 0 mV (Table 1). The unitary single-channel current to voltage (i - V) relation for Na^+ currents shows a slope conductance of 20 pS between -80 and -140 mV. The linear regression line extrapolates to a voltage intercept considerably negative to 0 mV, consistent with the inwardly-rectifying whole-cell current over the same voltage range. At negative potentials, unitary Ca^{2+} currents are much smaller than unitary Na^+ currents even though reversal potentials indicate that $P_{Ca}/P_{Na} > 1$. We interpret these findings in terms of Ca^{2+} binding sites in the pore, qualitatively similar to those in voltage-gated Ca^{2+} channels¹⁹. Accordingly, with 130 mM external Na^+ , addition of 1.5 mM Ca^{2+} to the external solution reduced the unitary current amplitude from 1.35 ± 0.07 pA to 1.08 ± 0.19 pA, (means $\pm$ s.e.m., $n = 3$). Addition of 100 μ M Ca^{2+} had little effect.

These experiments provide direct evidence that ATP-activated channels allow significant Ca^{2+} entry in either the absence or presence of Na^+ ions. In this respect, the channels in arterial smooth muscle differ from ATP-activated channels found in neurons¹⁷ or skeletal muscle²³. The ATP-activated channels provide a pathway for Ca^{2+} entry with biophysical and pharmacological properties quite different from voltage-gated L- or T-type

Ca^{2+} channels that have also been found in ear artery cells²⁰. With a three-fold preference for Ca^{2+} over Na^+ , the ATP-activated channels are much less selective than voltage-gated Ca^{2+} channels, which may exhibit $P_{Ca}/P_{monovalent}$ ratios of up to 10^4 , but an order of magnitude more selective than the nicotinic acetylcholine receptor channel²⁴. The ATP-activated channel resembles the NMDA-activated cation channels in showing mild selectivity for Ca^{2+} over Na^+ (ref. 25), but differs in its response to Mg^{2+} . Mg^{2+} ions readily permeate the ATP-activated channel but block NMDA channels in a potent and strongly voltage-dependent manner^{10,11}.

The open channel properties of the ATP-activated channels are consistent with the voltage-dependence of excitatory junction currents in submucous arterioles²⁶, and are thus appropriate to explain non-adrenergic excitatory junction potentials arising from sympathetic nerve stimulation²⁷. The ATP-activated channels can contribute to entry of activator Ca^{2+} by depolarizing the cell and promoting the opening of voltage-gated channels and by directly passing Ca^{2+} when Ca^{2+} -activated K^+ channels keep the membrane potential too negative to activate voltage-gated Ca^{2+} channels²⁷. This mechanism of Ca^{2+} entry may contribute to contractions evoked by sympathetic nerve stimulation even after block of α -adrenoceptors and dihydropyridine-sensitive channels²⁸. According to theory¹⁸, with 130 mM Na^+ and 1.8 mM Ca^{2+} outside and $P_{Ca}/P_{Na} \approx 3$, Ca^{2+} would carry about one-sixteenth of the ATP-activated inward current at -75 mV, or possibly more if ion-ion interactions were taken into account.

Our recordings from cell-attached and excised patches suggest a direct coupling between nucleotide binding and channel opening, without involvement of a readily diffusible messenger. In this regard, the ATP-activated Ca^{2+} entry in smooth muscle cells differs from Ca^{2+} entry promoted by f-Met-Leu-Phe in neutrophils, which is mediated by release of Ca^{2+} from intracellular stores¹², or from mitogen-activated Ca^{2+} -permeable channels in T lymphocytes, which are activated by inositol 1,4,5-trisphosphate¹³. It will be important to distinguish between direct and indirect mechanisms of activation of Ca^{2+} influx in liver, gland and endothelial cells and other systems where receptor-operated Ca channels have been postulated.

This work was supported by grants from USPHS, Marion Laboratories, The Miles Institute and SK&F Research (UK). We thank J. Isaacson for technical assistance, R. Y. Tsien for helpful comments and F. Sigworth for loan of equipment.

Received 8 April; accepted 22 May 1987.

- Neher, E. *Nature* **326**, 242 (1987).
- Bolton, T. B. *Physiol. Rev.* **59**, 607–718 (1979).
- Van Breemen, C., Aaronson, P. & Loutzenhizer, R. *Pharmac. Rev.* **30**, 167–208 (1979).
- Somlyo, A. P. & Somlyo, A. V. *J. Pharmac. exp. Ther.* **159**, 129–145 (1968).
- Reinhart, P. H., Taylor, W. M. & Bygrave, F. L. *Biochem. J.* **220**, 35–42 (1984).
- Putney, J. W. Jr *Pharmac. Rev.* **30**, 209–245 (1978).
- Exton, J. H. *Trends Pharmac. Sci.* **3**, 111–115 (1982).
- Hallam, T. J. & Rink, T. J. *FEBS Lett.* **186**, 175–179 (1985).
- Burnstock, G. & Kennedy, C. *Circulation Res.* **58**, 319–330 (1986).
- Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462–465 (1984).
- Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 261–263 (1984).
- von Tscherner, V., Prod'homme, B., Baggiolini, M. & Reuter, H. *Nature* **324**, 369–372 (1986).
- Kuno, M. & Gardner, P. *Nature* **326**, 301–304 (1987).
- Benham, C. D., Bolton, T. B., Byrne, N. G. & Large, W. A. *J. Physiol., Lond.* (in the press).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Kennedy, C. & Burnstock, G. *Blood Vessels* **22**, 145–155 (1985).
- Krishnal, O. A., Marchenko, S. M. & Pidoplichko, V. I. *Neurosci. Lett.* **35**, 41–45 (1983).
- Fatt, P. & Ginsborg, B. L. *J. Physiol., Lond.* **142**, 516–543 (1958).
- Tsien, R. W., Hess, P., McCleskey, E. W. & Rosenberg, R. L. *A. Rev. Biophys. Chem.* **16**, 265–290 (1987).
- Benham, C. D., Hess, P. & Tsien, R. W. *Circulation Res.* (in the press).
- Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752–754 (1981).
- Yellen, G. *Nature* **296**, 357–359 (1982).
- Kolb, H.-A. & Wakelam, M. J. O. *Nature* **303**, 621–623 (1983).
- Adams, D. J., Dwyer, D. M. & Hille, B. *J. gen. Physiol.* **75**, 493–510 (1980).
- MacDermott, A. B., Mayer, M. L., Westbrook, G. L., Smith, S. J. & Barker, J. L. *Nature* **321**, 519–522 (1986).
- Finkel, A. S., Hirst, G. D. & Van Helden, D. F. *J. Physiol., Lond.* **351**, 87–98 (1984).
- Bolton, T. B. & Large, W. A. *Q. J. Exp. Physiol.* **71**, 1–28 (1986).
- Suzuki, H. & Kou, K. *Jap. J. Physiol.* **33**, 743–756 (1983).
- Benham, C. D. & Bolton, T. B. *J. Physiol., Lond.* **381**, 385–406 (1986).